



Virginia Commonwealth University
VCU Scholars Compass

Theses and Dissertations

Graduate School

2015

A Novel Antimicrobial Drug Discovery Approach for the Periodontal Pathogen *Porphyromonas gingivalis*

Victoria N. Stone
Virginia Commonwealth University

Follow this and additional works at: <https://scholarscompass.vcu.edu/etd>



Part of the [Bacteriology Commons](#)

© The Author

Downloaded from

<https://scholarscompass.vcu.edu/etd/4008>

This Dissertation is brought to you for free and open access by the Graduate School at VCU Scholars Compass. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of VCU Scholars Compass. For more information, please contact libcompass@vcu.edu.

© Victoria N. Stone 2015
All Rights Reserved

A Novel Antimicrobial Drug Discovery Approach for the Periodontal Pathogen *Porphyromonas gingivalis*

A dissertation submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy at Virginia Commonwealth University

By

Victoria N. Stone

Bachelor of Science, Florida State University, Tallahassee, FL

Advisor: Ping Xu, Ph.D.

Professor, Philips Institute of Oral Health Research, School of Dentistry
Affiliate Associate Professor, Department of Microbiology and Immunology

Virginia Commonwealth University
Richmond, Virginia
October 2015

“What seems to us as bitter trials are often blessings in disguise”

-Oscar Wilde-

Acknowledgements

I would first like to express my sincere gratitude to my advisor, Dr. Ping Xu, for his support and guidance throughout this process. While it wasn't always easy, you allowed me build something from nothing. You provided me with a wonderful opportunity to learn and grow and for that I will be forever thankful.

Thanks to Dr. Xiuchun Ge for always taking time to answer my questions, even when you were busy.

Thanks to all the past and present members of my lab for sharing in this experience with me.

To all of the Philips Institute, especially Margaret Poland and Dung Pham. Wherever I go next, I hope there are people who are as caring and helpful as you both were.

I would also like to thank the members of my committee. Dr. Yan Zhang, thank you for taking the time to sit down and talk with me. I will always be grateful for your help in moving my project forward. Dr. Michael W. Holmes, thank you for listening to me when I felt lost or was overwhelmed. Dr. Todd Kitten thank you for your open door and allowing me to use absolutely everything in your lab, and Dr. Glen Kellogg thank you for your advice and providing the invaluable resources I needed for this project. Without all of your advice and guidance I would not be here today.

Thanks to Dr. Darrell Peterson for help during my project and allowing me to take over your office for my experiments.

Much love and thanks to my family, whose support gave me the strength to further my education.

I would also like to give a special thanks to Han Solo. You made me laugh, encouraged me and comforted me when I needed it. I'm glad that I met you.

I would like to acknowledge Virginia Commonwealth University for providing me with the opportunity to pursue a graduate degree.

Finally, thanks to the countless number of wonderful people I have met along the way. Your advice and conversations made the days easier.

I dedicate this magnum opus ☺ to my family, whom I love beyond all words. Daddy, you are everything I could ever hope to be; Mrs. Yates, thank you for always being there even when I didn't want to listen; Jr., Velencia and Veronica thanks for the calls, the gossip and the distractions.

Table of Contents

ACKNOWLEDGEMENTS	III
TABLE OF CONTENTS.....	V
LIST OF TABLES	VII
LIST OF FIGURES	VIII
LIST OF ABBREVIATIONS AND SYMBOLS.....	X
ABSTRACT	XIV
INTRODUCTION.....	1
I. THE ORAL MICROBIOME.....	1
Dental plaque.....	2
Periodontal disease.....	3
II. <i>PORPHYROMONAS GINGIVALIS</i>	7
Virulence factors.....	7
Keystone pathogen hypothesis.....	10
<i>P. gingivalis</i> as a potential target.....	11
III. ANTIBIOTICS	12
Antibiotic resistance.....	13
Antibiotics and the microbiome.....	15
Strategies in antimicrobial drug discovery.....	16
IV. SUMMARY	20
V. OUTLINE.....	21
CHAPTER ONE.....	23
BACKGROUND.....	24
MATERIALS AND METHODS	27
RESULTS.....	30
Defining our framework of essential gene prediction	30
Prediction of essential genes by gene annotation in <i>P. gingivalis</i> strain W83	32
Functional analysis and examination of predicted essential genes by biological categories.....	39
Cross-validation of prediction with experimental validation of essential genes in <i>P. gingivalis</i> strain W83	40
Comparison of <i>P. gingivalis</i> and <i>S. sanguinis</i> essential genes	47
Identification and assessment of <i>meso</i> -diaminopimelate as a target	47
DISCUSSION	57
CHAPTER TWO.....	66
BACKGROUND.....	67
MATERIALS AND METHODS	69
RESULTS.....	74

<i>In silico</i> analysis of <i>m</i> -Ddh structure and function.....	74
<i>In vitro</i> analysis of <i>m</i> -Ddh structure and function	77
Generation of pharmacophore model	88
High-throughput virtual screening for identification of small-molecule inhibitors. .	91
Molecular docking and scoring of small-molecule inhibitors	91
<i>In silico</i> analysis of binding confirmations.....	92
DISCUSSION	97
CHAPTER THREE	101
BACKGROUND.....	102
MATERIALS AND METHODS	104
RESULTS.....	107
Evaluation of enzymatic inhibition against <i>m</i> -Ddh	107
Evaluation of cellular inhibition against <i>m</i> -Ddh.....	109
SAR evaluation of compound 4 core scaffold	111
Evaluation of inhibition mechanism.....	117
DISCUSSION	122
CONCLUSION	127
APPENDIX	131
LITERATURE CITED	144
VITAE	163

List of Tables

Table 1. Non-annotated predicted essential genes in <i>P. gingivalis</i> strain W83.	38
Table 2. Predicted <i>P. gingivalis</i> -selective essential genes.	48
Table 3. Kinetic analysis of <i>m</i> -Ddh mutants.	85
Table 4. 2D structures and scoring of top-ranking compounds	93
Table 5. Analysis of <i>m</i> -Ddh enzymatic inhibition with active compounds.	108
Table 6. Analysis of whole-cell inhibition with active inhibitors	110
Table 7. SAR of compound 4 analogues.	116
Supplemental Table 1. Predicted putative essential genes in <i>P. gingivalis</i> strain W83.	131

List of Figures

Figure 1. Workflow for two methods of antimicrobial drug discovery.....	17
Figure 2. Workflow schematic for the prediction model framework.	33
Figure 3. Predicted essential genes and corresponding pathways for <i>P. gingivalis</i> strain W83...	35
Figure 4. Pathway association and COG distribution for predicted essential genes in <i>P. gingivalis</i> strain W83.	41
Figure 5. Cross-validation of prediction to experimental essential gene data.	45
Figure 6. Variant pathways of lysine and <i>meso</i> -diaminopimelate biosynthesis.	52
Figure 7. Allelic replacement mutagenesis for predicted essential gene target, <i>m</i> -Ddh.....	55
Figure 8. Sequence analysis of <i>m</i> -Ddh.	58
Figure 9. 3D structural analysis of <i>m</i> -Ddh from <i>P. gingivalis</i>	60
Figure 10. Schematic for the <i>m</i> -Ddh catalyzed biochemical reaction.	75
Figure 11. Structural domains of <i>m</i> -Ddh from <i>P. gingivalis</i>	78
Figure 12. Molecular docking analysis of <i>m</i> -DAP binding model.	80
Figure 13. Kinetic analysis of purified <i>m</i> -Ddh from <i>P. gingivalis</i> strain W83.....	83
Figure 14. Site-directed mutagenesis of binding pocket.	86
Figure 15. Generation of pharmacophore model.	89
Figure 16. Optimal binding mode for top-ranking compounds.	95
Figure 17. Time-kill analysis of compound treated <i>P. gingivalis</i>	112
Figure 18. SEM analysis of compound treated <i>P. gingivalis</i> cells.	114
Figure 19. 3D analysis of binding model for active inhibitors.	118
Figure 20. Inhibition mechanism of active compounds in regards to substrate, <i>m</i> -DAP and co-substrate, NADP ⁺	120
Supplemental Figure 1. Example of transposon insertion for <i>P. gingivalis</i> strain W83.	140

Supplemental Figure 2. SDS PAGE and gel filtration analysis of purified <i>m</i> -Ddh from <i>P. gingivalis</i>	142
--	-----

List of abbreviations and symbols

<	Less than
≈	Approximately
2D	Two Dimensional
3D	Three Dimensional
α	Alpha
Å	Ångstrom
A	Alanine
a.a.	Amino acid
Arg	Arginine
Asp	Aspartate
β	Beta
<i>B. subtilis</i>	<i>Bacillus subtilis</i>
bp	Base pairs
°C	degrees Celsius
<i>C. difficile</i>	<i>Clostridium difficile</i>
<i>C. glutamicum</i>	<i>Corynebacterium glutamicum</i>
CBDD	Computer-based Drug Discovery
CD	Cluster of Differentiation
CDC	Center for Disease Control
CFU/ml	Colony Forming Units per milliliter
COG	Cluster of Orthologous Groups
D	Aspartate
Da	Daltons
DEG	Database of Essential Genes
DMSO	Dimethyl Sulfoxide
ε	Epsilon

E	Glutamate
<i>E. coli</i>	<i>Escherichia coli</i>
Erm	Erythromycin
Erm ^R	Erythromycin Resistance
Gly	Glycine
GOLD	Genetic Optimization for Ligand Docking
GSK	GlaxoSmithKline
HINT	Hydropathic INteractions
His	Histidine
h	Hours
<i>H. influenza</i>	<i>Haemophilus influenza</i>
HOMD	Human Oral Microbiome Database
HPLC	High Performance Liquid Chromatography
HTS	High-throughput Screening
IC ₅₀	Half Maximal Inhibitory Concentration
IL	Interleukin
K	Lysine
kDa	Kilodaltons
KEGG	Kyoto Encyclopedia of Genes and Genomes
Kgp	Lysine Specific Gingipain
LPS	Lipopolysaccharide
mM	Millimolar
MBC	Minimum Bactericidal Concentration
Met	Methionine
MIC	Minimal Inhibitory Concentration
<i>m</i> -DAP	<i>meso</i> -diaminopimelate
<i>m</i> -Ddh	<i>meso</i> -diaminopimelate dehydrogenase
mg	Milligrams

min	Minutes
ml	Milliliters
MoA	Mechanism of Action
NADP(H)	Nicotinamide adenine dinucleotide phosphate (reduced)
NIH	National Institute of Health
nm	Nanometers
ND	Not Determined
OD	Optical Density
<i>P. gingivalis</i>	<i>Porphyromonas gingivalis</i>
<i>P. intermedia</i>	<i>Prevotella intermedia</i>
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PMN	Polymorphonuclear
PDB	Protein Data Bank
Phe	Phenylalanine
R	Arginine
Rgp	Arginine Specific Gingipain
<i>S. aureus</i>	<i>Staphylococcus aureus</i>
<i>S. pneumoniae</i>	<i>Streptococcus pneumoniae</i>
<i>S. sanguinis</i>	<i>Streptococcus sanguinis</i>
SAR	Structure Activity Relationship
SBVS	Structure-based Virtual Screening
S.D.	Standard Deviation
sec	Seconds
Thr	Threonine
TLR	Toll-like Receptor
TNF	Tumor Necrosis Factor
Trp	Tryptophan

TSB	Tryptic Soy Broth
Tyr	Tyrosine
μl	Microliters
μM	Micromolar
WT	Wild Type
Y	Tyrosine

Abstract

A Novel Antimicrobial Drug Discovery Approach for the Periodontal Pathogen *Porphyromonas gingivalis*

By: Victoria N. Stone

A dissertation submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy at Virginia Commonwealth University

Virginia Commonwealth University, 2015

Advisor: Ping Xu, Ph.D.

Professor, Philips Institute of Oral Health Research, School of Dentistry
Affiliate Associate Professor, Department of Microbiology and Immunology

The human body is colonized by more than 100 trillion microbes which make up an essential part of the body and plays a significant role in health. We now know the over use and misuse of broad-spectrum antibiotics can disrupt this microbiome contributing to the onset of disease and runs the risk of promoting antibiotic resistance. With antibiotic research still on the decline, new strategies are greatly needed to combat emerging pathogens while maintaining a healthy microbiome. We therefore set out to present a novel species-selective antimicrobial drug discovery strategy.

Disruption of the homeostasis within the oral cavity can trigger the onset of one of the most common bacterial infections, periodontal disease. Even though the oral cavity is one of the most diverse sites on the human body, the Gram-negative colonizer, *Porphyromonas gingivalis* has long been considered a key player in the initiation of periodontitis, suggesting the potential for novel narrow-spectrum therapeutics. By targeting key pathogens, it may be possible to treat periodontitis while allowing for the

recolonization of the beneficial, healthy flora. Therefore, we set out to use *P. gingivalis* and periodontal disease as a model for pathogen-specific antimicrobial drug discovery. In this study we present a unique approach to predict essential gene targets selective for the periodontal pathogen within the oral environment. Using our knowledge of metabolic networks and essential genes we identified a “druggable” essential target, *meso*-diaminopimelate dehydrogenase, which is found in a limited number of species. This enzyme, *meso*-diaminopimelate dehydrogenase from *P. gingivalis*, was first expressed and purified, then characterized for enzymatic inhibitor screening studies. We then applied a computer-based drug discovery method, combining pharmacophore models, high-throughput virtual screening and molecular docking. Utilizing the ZINC database we virtually screened over 9 million small-molecules to identify several potential target-specific inhibitors. Finally, we used target-based and whole-cell based biochemical screening to assess *in vitro* activity. We conclude that the establishment of this target and screening strategy provides a framework for the future development of new antimicrobials and drug discovery.

Introduction

I. The Oral Microbiome

The term “microbiome” was coined by Joshua Lederberg describing “the ecological community of commensal, symbiotic, and pathogenic microorganisms that literally share our body space” [1, 2]. The start of the microbiome era resulted in the question of how commensal microorganisms contribute to human health and disease and sparked new areas of research. This directly led to the NIH-sponsored Human Microbiome Project. We now know that microbes outnumber our cells 10 to 1, making bacteria an important component of the human body. Through this project it was discovered the oral cavity is one of the most complex and diverse sites [3], containing up to 1000 phylotypes composed of viruses, protozoa, fungi, archaea and bacteria. However, since many oral bacteria have yet to be cultured, the number may be even higher [1, 4]. While the functional role of oral colonizers is not completely understood, studies indicate this area of the microbiome plays a role in maintaining oral health. For one, the occupation by commensal oral bacteria prevent the colonization of pathogens, a phenomenon known as colonization resistance [5]. Commensal bacteria occupy the niche, limiting the available space and preventing the establishment of foreign colonizers. As an indirect mechanism of colonization resistance, commensal oral colonizers can produce antagonistic substances against pathogenic species [4, 6]. For example, many streptococcal species can synthesize inhibitory substances that prevent colonization of other species. *Streptococcus sanguinis* produces hydrogen peroxide, which can inhibit the growth of methicillin-resistant *Staphylococcus aureus* [7] and *Streptococcus mutans*, a major contributor to dental caries [8]. Interestingly, studies also show the oral microbiome plays a functional role in systemic health. Inorganic nitrate obtained through the diet is reduced to

nitrite within the oral cavity, absorbed into the bloodstream and then converted to nitric oxide. This is beneficial, as metabolic studies show nitric oxide is important for maintaining cardiovascular health through improved mitochondrial function and reduced blood pressure [9]. Petersson *et al.* linked oral bacteria to this metabolic role by noting the use of antimicrobial mouthwash attenuated the nitrate reduction and abolished the associated decrease in blood pressure [10].

Dental plaque. How is the complex oral microbiome established? The bacteria have the ability to colonize the hard and soft surfaces (e.g., teeth, surfaces of the tongue, and epithelium) of the oral cavity to form biofilms, more commonly known as dental plaque. Oral biofilms are polymicrobial 3D structures composed of proteins, polysaccharides, nucleic acids, salts and cells allowing for microbial interactions and cell-cell communication. It also provides the population with nutrients, protection from antibiotics and other environmental threats [11, 12]. The development of dental plaque is a systematic process occurring through five stages [12]. The first stage begins almost immediately following brushing. The tooth surface obtains a coating of glycoprotein deposits derived from saliva called the acquired pellicle. Through interactions with cell surface structures, the pellicle mediates the initial bacterial attachment. It can be 1-20 layers thick and within a clinically healthy cavity is composed of Gram-positive cocci, mainly of the streptococcus species [13, 14]. These early colonizers are aerobes or facultative anaerobes as they are able to tolerate the higher oxygen levels present [14, 15]. The second stage involves a stronger attachment of the bacteria to the tooth surface. Following initial attachment to the pellicle, weak, long-ranged interactions are formed that allows the bacteria to reverse the attachment to the pellicle and tooth surface. Stronger short-range, irreversible attachments between structures on the bacterial cell surface and the complementary receptors on the pellicle

are then formed [15]. The third step involves adhesion and aggregation of later colonizers to previously attached early colonizers. Inter-bacterial adhesion-receptor interactions are initiated, resulting in the diversification of the biofilm [13], as well as generating synergistic and antagonistic interactions between species [15]. Fourth is the multiplication of the colonized bacteria; this step involves cell division of colonized bacteria and expansion of the biofilm into a complex 3D environment. Polymer composed of soluble and insoluble glucans is produced forming the extracellular matrix. This matrix provides structural integrity and retains nutrients within the biofilm. At this stage the biofilm is around 100-300 layers thick, generating an oxygen depleted environment that allows for an increase in the number of Gram-negative anaerobic bacteria [14, 16]. Last is active detachment. In response to environmental cues, the biofilm can detach from the surface, disperse and colonize at different locations [15].

Periodontal disease. Typically, with proper oral hygiene, the oral microbiome exists in a beneficial or benign state. However, ecological changes in the oral cavity can alter the environment, shifting the dominant species from Gram-positive facultative anaerobes to Gram-negative anaerobes. The resulting change in the environment can lead to a major health concern within the oral cavity known as periodontal disease [12]. Periodontal disease refers to a group of chronic inflammatory processes that affect the gingival tissue and surrounding tooth supporting structures [17]. Periodontal disease is a progressive disease that occurs in distinct stages.

Gingivitis, the most common form, is the early non-destructive stage, which is composed of the initial, early and established lesion stages. It is reversible with proper cleaning and maintenance; however, without treatment gingivitis may progress into a more chronic infection. The advanced lesion stage marks the transition from gingivitis to periodontitis, the non-reversible destructive form. Periodontitis is characterized by severe inflammation, plaque calcification, bone

resorption, and tooth loss [18]. During the early lesion stages of the infection, within 2-10 days, accumulation of dental plaque triggers the host innate immune response where resident mast cells produce histamine causing the endothelium to release IL-8 [19, 20]. Additionally, macrophages recognize pathogen-associated molecular patterns (i.e., peptidoglycan, lipopolysaccharide and foreign DNA) through toll-like receptors (TLRs), triggering the uptake of the bacteria through phagocytosis and the release of cytokines such as IL-1 β , IL-6 and TNF- α . The release of these pro-inflammatory cytokines cause vascular dilation of blood vessels, increasing permeability and blood flow. This also leads to the expression of adhesion molecules on the endothelium surface, aiding in the recruitment and sticking of polymorphonuclear neutrophils (PMNs) and activation of complement [21]. PMNs migrate to the epithelium and gingival crevice releasing more IL-8, amplifying inflammation. However, the PMNs are unable to phagocytose the bacteria, which are now associated within a thick biofilm. After approximately 2-3 weeks of biofilm accumulation, antigen-presenting cells begin to stimulate naïve T-cells. At this stage if T-cells have not eliminated the bacteria, an established lesion is set. This stage is defined by B-cell activation to plasma cells [20, 21]. Th₂ cells activate B-cells and conversion of plasma cells leads to the secretion of antibodies. CD4⁺ T-cells continue to secrete cytokines contributing to collagen degradation and inflammation [21]. At this point, if the infection is not cleared, gingivitis progresses into periodontitis, corresponding to the advanced lesion stage and marked by severe collagen degradation. The formation of deep periodontal pockets allows for the migration of the biofilm, increased production of pro-inflammatory cytokines and continued tissue degradation [21]. Plasma cells are now the dominant immune cells. Continued T-cell response leads to the release of interleukins, TNF- α and prostaglandin E₂, being mainly responsible for the alveolar bone destruction [22].

Clinical Significance. The high prevalence of periodontal disease in developing and industrialized countries has made it a significant public health concern [14, 23]. The economic burden of the disease and subsequent treatment can be overwhelming. Oral related issues are the fourth most expensive disease, contributing up to 5-10% of health-care related costs [23, 24]. There are also corresponding psychological effects as the deterioration of the gum line and tooth loss can lead to embarrassment and low self-esteem [25, 26]. In addition, poor oral hygiene has been associated with systemic comorbidities. Brushing and invasive dental procedures allow bacteria to enter the bloodstream and disseminate to other sites, such as the brain, lungs and liver, linking periodontal disease to pregnancy complications, respiratory, cardiovascular and cerebrovascular diseases [23, 27, 28].

Epidemiology. Periodontal diseases are some of the most common bacterial infections worldwide. It is estimated that approximately 82% of US children and 50% of adults display overt signs of gingivitis with gingival inflammation and bleeding. This percentage is higher for both children and adults in developing countries [29]. For the more severe form, it is estimated that around 46% or 64.7 million US adults 30 years and older have some stage of periodontitis. Unfortunately, this only increases with age. Chronic periodontitis is most prevalent in adults or seniors 65 years and older, with approximately 70.1% of the population afflicted [30]. Epidemiological studies have shown clear disparities in the estimated populations affected by periodontal disease. There are higher incidences of periodontitis in men than women (56.4% vs 38.4%) and higher frequencies in individuals of African (59.1%) and Hispanic descent (63.5%) [30]. There is also a socioeconomic trend, with individuals with lower socioeconomic status more likely to present with periodontitis than those with a higher socioeconomic status [29, 31]. Specifically, approximately 65.4% of individuals in the US living below the federal poverty line

and 66.9% of individuals with less than a high school education have reported cases of periodontitis associated tooth loss.

Risk Factors. A risk factor is defined as a variable that has been associated with increased risk of disease or infection [32] and studies show certain elements contribute to the development of periodontitis. Smoking, diabetes mellitus, age and stress are all strongly linked to the development of the disease [32, 33]. While these factors may not directly contribute to the disease, it is likely that the systemic changes (chronic inflammation and altered immune response) associated with these factors disrupt the oral ecosystem and contribute to the onset and prevalence of the disease. For example, patients with diabetes exhibit increased vascular permeability and impaired neutrophil migration contributing to their increased susceptibility for the disease [34]. However, it is important to note inherent non-modifiable factors such as genetic susceptibility and host response play a role in the progression of periodontal disease, as some people may exhibit one or more of the associated risk factors and never develop the disease [32].

It is widely accepted that the microbial population is a significant risk factor. Initially, it was believed periodontal disease was a result of the overwhelming presence of total bacteria in what was known as the non-specific plaque hypothesis [35-37]. However, in the 1970s, the specific plaque hypothesis was offered, stating that certain species played an etiological role in the onset. Evidence for this theory began with cariogenic bacteria. Loesche *et al.* noticed that antimicrobial therapy with kanamycin targeting oral streptococci reduced caries formation [35, 38, 39]. Following this, Socransky *et al.* offered supporting evidence by studying the complex poly-microbial environment based on the bacterial relationships corresponding to health and disease state. Socransky recognized that late colonizers *Porphyromonas gingivalis*, *Treponema denticola* and *Tannerella forsythia* (formerly *Bacteroides forsythus*) were isolated together from

diseased sites and expressed numerous virulence factors, corresponding to the effects of periodontitis. This bacteria group makes up what is still known as the ‘red complex’ [40, 41].

II. *Porphyromonas gingivalis*

One of the main pathogens within the oral cavity is *Porphyromonas gingivalis*. *P. gingivalis* is a Gram-negative non-motile, rod-shaped organism that belongs to the phylum *Bacteroidetes* and has long been considered a major causative agent in the onset and progression of periodontal disease [42, 43]. Identifying characteristics of the bacterium include anaerobic respiration, which is the degradation of amino acids for energy and the accumulation of hemin on the cell surface for the acquisition of iron, forming the characteristic black colonies on blood agar. *P. gingivalis* has been isolated from the oral cavity [18], the respiratory tract [44, 45], and from the vagina of women with bacterial vaginosis [46]. The genome of virulent *P. gingivalis* strain W83 was sequenced in 2003 by Nelson *et al.* revealing a 2.3 Mbp genome with 2,015 protein-coding genes [43]. *P. gingivalis* is a secondary colonizer thriving in nutrient rich areas with reduced oxygen levels, attaching to streptococci early colonizers [18, 47] as well as other late colonizers such as *Fusobacterium nucleatum* and *Treponema denticola* [43, 48]. The bacterium is a part of the “red complex” as it is readily isolated from diseased sites and commonly found with other known oral pathogens [40, 49]. There is a positive correlation between an increase in periodontal pocket depth and *P. gingivalis* colonization [50]. Additionally, studies show there is improvement in periodontal health following the reduction of *P. gingivalis* [51]. Colonization and virulence of *P. gingivalis* is aided by the expression of protein adhesins, proteinases and hemagglutinins [18].

Virulence factors. A virulence factor is defined as a molecule produced by a pathogenic organism that contributes to the destructive nature of a disease and enables the pathogen to

survive within a specific host [52]. As previously mentioned, numerous studies have revealed a plethora of virulence factors that contribute to the colonization of *P. gingivalis*, the subversion of the host immune response and the destruction of the periodontal structures [18, 52].

Gingipains. *P. gingivalis* possesses several proteases; however, the most studied belongs to a group of cell associated and secreted cysteine proteases known as gingipains. Gingipains contribute to a variety of virulent functions and have been classified based on their enzymatic specificity: arginine specific gingipains, Rgp and lysine specific gingipains, Kgp [18, 52]. One major purpose of gingipains is nutrient acquisition. Expression of Rgp and Kgp can induce epithelial cell death and tissue degradation that results in attainable nutrients for the bacteria within the environment [53]. It is also thought that *P. gingivalis* utilizes gingipains to break down transferrin and hemoglobin to obtain the iron necessary for survival [54]. This theory was supported by studying gingipain-deficient *P. gingivalis* cells. These mutants lost their characteristic black pigment on blood agar and their ability to bind hemoglobin [55, 56]. Gingipains also play a role in the deregulation of the host immune defense. Studies show that gingipains possess complement-like convertase activity by cleaving complement components and allowing for immune dysregulation [57, 58].

Fimbriae. The fimbriae are long protein structures peritrichously arranged on the outer cell surface of the bacteria and facilitate adhesion to the surfaces of the oral cavity as well as other bacteria. Addition of purified monoclonal anti-fimbriae antibodies prevented binding and colonization of *P. gingivalis* to oral epithelial cells [59, 60] and strains of *P. gingivalis* with short or very few fimbriae were typically non-adherent to host cells [52]. The fimbriae may also play a role in periodontitis associated tissue destruction as immunization against *P. gingivalis* fimbriae in a rat model prevented tissue breakdown and characteristic periodontal destruction [52, 61].

Hemagglutinins. Hemagglutinins are microbial lectins or adhesins expressed on the bacterial cell surface and are essential for the initial attachment and establishment of infection. Binding is mediated through glycan receptors on host cells which may be facilitated by the fimbriae. Consequently, studies show antibodies against hemagglutinin HagB in *P. gingivalis* significantly reduced attachment to endothelial cells [62]. Hemagglutinins have long been associated with virulence as they may agglutinate erythrocytes [63] and sequence data shows that certain hemagglutinins can be co-expressed with genes containing proteolytic activity for the acquisition of iron [17]. The ability to block attachment and therefore colonization have led to new studies into vaccines targeting hemagglutinins as potential therapeutics [63].

Lipopolysaccharide. Lipopolysaccharide (LPS) is a glycolipid found on the outer membrane of the bacterial cell surface. It is a potent immunostimulant, generating an array of immunological responses triggering inflammation and tissue destruction. *P. gingivalis* LPS has been shown to be endotoxic since the bacteria has the ability to release LPS in vesicles [64], which can then enter the host tissues [65, 66]. The interaction of LPS with the host tissues induces hallmark pro-inflammatory cytokines such as IL-1 β , TNF- α and IL-6. Recognition of LPS can directly stimulate osteoclasts leading to the release of IL-1 β and TNF- α from various immune cells causing the hallmark signs of periodontitis: bone resorption and tissue destruction [17]. Toll-like receptor 4 (TLR4) is the major signal transducer for bacterial lipopolysaccharides; however, the *P. gingivalis* lipid A moiety of LPS has been described as atypical. It is less potent compared to other bacterial LPS. Furthermore, it is agonistic against TLR2, while being both antagonistic and agonistic for TLR4 [67]. This may be due to the differences in the chemical composition in response to environmental cues (i.e., fatty acid chain heterogeneity [68], changes in acylation patterns or phosphorylation of the lipid A moiety [17, 67]). The ability to alter the

LPS structure can be beneficial to *P. gingivalis* survival through manipulation of the innate immune response. This response suppresses inflammation and immune clearance early during infection, then later induces inflammation for nutrient acquisition. Furthermore, studies show that the atypical *P. gingivalis* LPS is able to down-regulate the stimulatory effect of other bacterial LPS by TLR4, which may stimulate E-selection on endothelial cells, thereby preventing neutrophil migration and evading clearance [17, 69].

Keystone pathogen hypothesis. The ecological application of the term “keystone” refers to a species that is present in low numbers relative to the total population but still plays a major role in maintaining the integrity of the community [70]. Historically, *P. gingivalis* is known to be highly associated with the clinical signs of periodontal disease such as oral lesions and bone loss, while accounting for a small fraction of the oral population [71]. Evidence suggesting *P. gingivalis* could be a “keystone pathogen” began from a study when *P. gingivalis* colonized at less than 0.01% in specific-pathogen free mice, induced destructive periodontitis. These effects were not observed in germ-free mice, indicating the onset of periodontal disease is related to changes in the microbial population induced by *P. gingivalis* [58]. This was further characterized by showing that the colonization of *P. gingivalis* selectively modified the host immune response leading to the disruption of the normal host microbiota [58]. Following colonization, *P. gingivalis* prevented epithelial cell secretion of IL-8, a chemoattractant for PMN that hindered neutrophil migration and attenuated the first line of immune defense [58, 72, 73]. Interestingly, this colonization also prevented complement activation, which facilitated in the survival of the periodontal pathogen, the increase in commensal bacterial load and induction of bone loss [58]. *P. gingivalis* cleaves the complement component C5 into C5a by gingipains [74] and the downstream effect of C5aR in conjunction with TLR2 stimulated by the pathogen-associated

LPS leads to a break within the signaling pathway preventing phagocytosis. This interaction is important to the dysregulation as gingipain-deficient strains of *P. gingivalis* show no changes in the oral micro-environment maintaining the bacterial balance [75, 76]. The interaction with TLR2 prevents the production of IL-12 cytokines, but increases proinflammatory mediators IL-1 β , IL-6 and TNF- α . This selective immunosuppressive mechanism aids *P. gingivalis* in resisting immune clearance while allowing for cytokines that are involved in bone resorption and tissue degradation to generate a nutrient rich, inflamed site [77]. The now altered oral environment may select for asaccharolytic organisms such as *P. gingivalis* since they are able to utilize available nutrients. In addition, the degeneration of alveolar bone provides new niches for pathogenic species. *P. gingivalis* also alters macrophage function. The bacteria interact with lipid rafts on the macrophage cell surface, allowing for a close interaction between FimA and CXCR4 and TLR2 on the macrophage cell surface. Crosstalk between the two receptors increases cAMP levels and inhibits the oxidative burst facilitated clearance by macrophages [45, 78].

***P. gingivalis* as a potential target.** Corresponding to the evidence of being a keystone pathogen, studies of immunization directed specifically against *P. gingivalis* have shown promise as a therapeutic strategy for periodontal disease. Mice immunized with purified *P. gingivalis* capsular polysaccharide developed high IgM and IgG serum titers, which elicited protection against *P. gingivalis*-induced bone loss [79]. Several studies have shown success in targeting outer-membrane proteins (OMP) [80-83]. For instance, *P. gingivalis*-challenged mice exhibited decreased oral lesions when immunized against recombinant forms of an OMP-like protein [83]. Another study using a rat model examining immunization against the adhesin motif of hemagglutinin HagB showed a reduction in periodontitis associated bone loss [62, 84, 85]. A study in non-human primates, vaccinating against an outer membrane protein or a lysine

gingipain protease complex significantly reduced alveolar bone loss and decreased *P. gingivalis* cell viability [86, 87]. Finally, a human clinical study providing repeated applications of monoclonal antibodies specific to a *P. gingivalis* protease complex prevented recolonization of the pathogen for approximately 9 months, leading to significant improvement of oral health [84, 88].

III. Antibiotics

Antibiotics or antimicrobials are some of the most commonly prescribed medications worldwide [89] and describe any molecule that targets essential functions (e.g., DNA replication, protein synthesis, cell wall synthesis and cell membrane) to kill or stop the growth of infectious organisms [90-92]. Around 1930, the modern “golden age” of antimicrobial research began, representing one of the greatest medical achievements [93, 94]. Beginning in the early 1900s, Paul Ehrlich hypothesized the possibility of selectively targeting pathogenic microbes while sparing the host [90]. However, it wasn’t until 1928 that the theory became a reality when Alexander Flemming made a fortuitous discovery that resulted in the wide spread use of penicillin [90]. Spanning a 40-year period, researchers and pharmaceutical companies examined microbial and fungal metabolites for naturally occurring antibiotics leading to discovery of the major antibiotic class scaffolds: cephalosporins, penicillins, quinolones and macrolides, still used today [90, 95]. During this period, steady progress was made with improvements in the form of second and third generation synthetic derivatives, resulting in various β -lactams, sulfonamides, and aminoglycosides then tetracyclines, macrolides and glycopeptides [90, 95, 96]. By the 1980s, development began to slow. The ideal natural sources had been exhausted, leading to a drop in progress and interest. Any newly designed antimicrobials, by that point, were chemical

modifications of common core scaffolds and targets, incrementally improving potency and efficacy. A few novel compounds were identified however, but none made it to market until years later. This included linezolid, an oxazolidinone class discovered in 1978, approved in 2000, daptomycin, a lipopeptide antibiotic discovered in 1986, approved in 2003 and retapamulin, a pleuromutilin approved in 2007 [93].

Antibiotic resistance. The discovery of antimicrobials brought about a significant turn in the constant battle against bacterial infections, increasing both the quality and span of the average life [92, 97]. However, the rise of antibiotic resistance has threatened to diminish this feat, rendering many antibiotics ineffective. Resistance is defined as a change in the bacteria's level of susceptibility to a drug, in regards to the normal population. This change can occur through mutations or the acquirement of heterologous genes coding for resistance by means of horizontal gene transfer [98]. Once expressed, the acquired gene provides a selective advantage for the pathogen when exposed to antibiotics leading to the spread and dominance of resistant strains within the population [90]. Mechanistically, bacterial resistance can occur in four ways. First, the modification of the drug target. For example, resistance to the glycopeptide vancomycin is mediated through the acquired *vanA* gene. VanA codes for an enzyme that synthesizes a peptidoglycan with one less hydrogen bond, decreasing vancomycin's affinity by 1000-fold. Second is the inactivation of the drug. Aminoglycoside antibiotics such as kanamycin are left ineffective due to the production of aminoglycoside-modifying enzymes that covalently attach a moiety (e.g., phosphate, acetyl or adenyl) to a key functional group of the antibiotic. This modification is a steric hinderance to binding of the ribosome. The third method is impermeability, such as the down-regulation of porins. Fourth, resistance can occur through the

expulsion of the drug. Streptogramin resistance by *E. faecalis* is partially due to the presence of ABC transporter Lsa, which removes the drug and prevents access to the target [98-101].

Antibiotic resistance not only increases healthcare related cost forcing the use of more expensive drugs and prolonging treatment times, but it is a significant threat to our morbidity and mortality [98]. By 2004, it was documented that approximately 70% of pathogenic bacteria were resistant to at least one antibiotic. This has resulted in at least two million illnesses and 23,000 deaths, adding an additional \$20 billion to healthcare cost [102]. Rising resistance in *Staphylococcus aureus* is one well-known example. This Gram-positive pathogen asymptotically colonizes 30% of the population and is known to be extremely adaptable against antibiotic pressure [103, 104]. By the beginning of 1950, shortly following the introduction of penicillin into clinical practice, a large proportion of hospital staphylococcal infections were already penicillin-resistant [105]. Initially only found throughout hospitals, within a few years, penicillin-resistant strains began causing community acquired infections. This led to the use of methicillin as the first drug of choice. Nevertheless, by 1961, cases of methicillin-resistant *Staphylococcus aureus* (MRSA) were on the rise in the United Kingdom and by the 1980s, MRSA infections were widespread within the United States [104, 106, 107]. From that point, vancomycin became the last line of effective antibiotics. Again, antibiotic pressure led to the emergence of vancomycin-intermediate *S. aureus* strains (VISA) and by 2002, cases of strains displaying complete resistance or vancomycin-resistant *S. aureus* strains (VRSA) were found within the US [104]. In 2006 the CDC reported antibiotic resistant *S. aureus* infections such as pneumonia, sepsis and necrotizing fasciitis, resulted in approximately 19,000 deaths [95, 102]. Re-emerging pathogens such as *Mycobacterium tuberculosis*, the etiological agent for tuberculosis (TB), add another layer of concern. While TB infections were once on the decline,

in 2013 there were around 480,000 new multi-drug resistant TB (MDR-TB) cases and 100 countries reporting cases of extensively-drug resistant TB (XDR-TB) strains resistant to first-line and second-line therapies [92, 108, 109]. XDR-TB is extremely difficult to treat, requiring at least two years of antimicrobial therapy. Additionally, effective therapy typically relies on antibiotic drugs that are expensive, toxic and potentially lethal [95].

Antibiotics and the microbiome. Humans are colonized with over 100 trillion bacteria, which led to a symbiotic host-microbe relationship [110]. The microbiome plays important functional roles, from establishing our immune system after birth [111, 112] to influencing how our brains process information [113, 114]. Since the Human Microbiome Project, there has been growing interest in the healthy human microbiome and how changes in it relate to the development of disease [3, 110, 115, 116]. Therefore, it is not surprising that disruption in that normal microbiota (as a result of the overuse or misuse of broad-spectrum antibiotics) has been linked to a multitude of systemic diseases. Reoccurring *Clostridium difficile* infection is a distinctive case of changes in the microbiome due to antibiotic therapy. Usage of antimicrobials disrupts the normal gut microbiota allowing for the colonization of *C. difficile*, resulting in fever, severe diarrhea and colitis [117]. Research has begun linking the alteration of the microbiome to obesity. Ley *et al.* noted significant changes in the dominant gastrointestinal microbial species in obese mice compared to lean mice [118]. Correspondingly, mice given the fecal extract from an obese twin subsequently led to weight gain, while the extract from the leaner twin did not [119]. Infants treated with ampicillin/gentamicin soon after birth showed a decrease in beneficial gut species that did not recover within 8 weeks post treatment and increased the incidence of weight gain [120, 121].

This brings in to play the rationale for alternative strategies such as narrow-spectrum or pathogenic-specific antimicrobials. Improvements in diagnostics have decreased the time it takes to identify the etiological agent and narrow-spectrum treatment can be more cost effective long-term. Antibiotics drug design against a limited population may also lead to faster discovery rates as targets will not need to be conserved or effective across a range of disparate species. With rising resistance, the necessity of maintaining the healthy microbial population and progress in general antimicrobial development of novel approaches are imperative.

Strategies in antimicrobial drug discovery. Drug discovery is a multi-step process spanning from target selection to drug optimization and clinical trials [122]. Historically, there have been two main approaches to antimicrobial drug discovery: empirical screening and target-based discovery (Fig 1). The empirical approach involved whole-cell based, trial and error screening of naturally occurring compounds or their synthetic derivatives. Compounds displaying growth inhibition against the desired spectrum of bacteria would be screened for eukaryotic cytotoxicity, followed by secondary screening to identify the mode of action. Biological assays would be used to verify the activity of the compound, ending with optimization and development for clinical trials [97]. By the early 1990s it was clear, even though successful in the past, this could no longer provide the results needed. The decreasing success of this process was due to the fact it was: (1) expensive and time consuming, while yielding a poor return for the effort [94, 123]; (2) had been depleted of easily obtainable products isolated from organisms culturable at the time [97]; and (3) changes in regulatory policies made it increasingly difficult to obtain clinical approval [124].

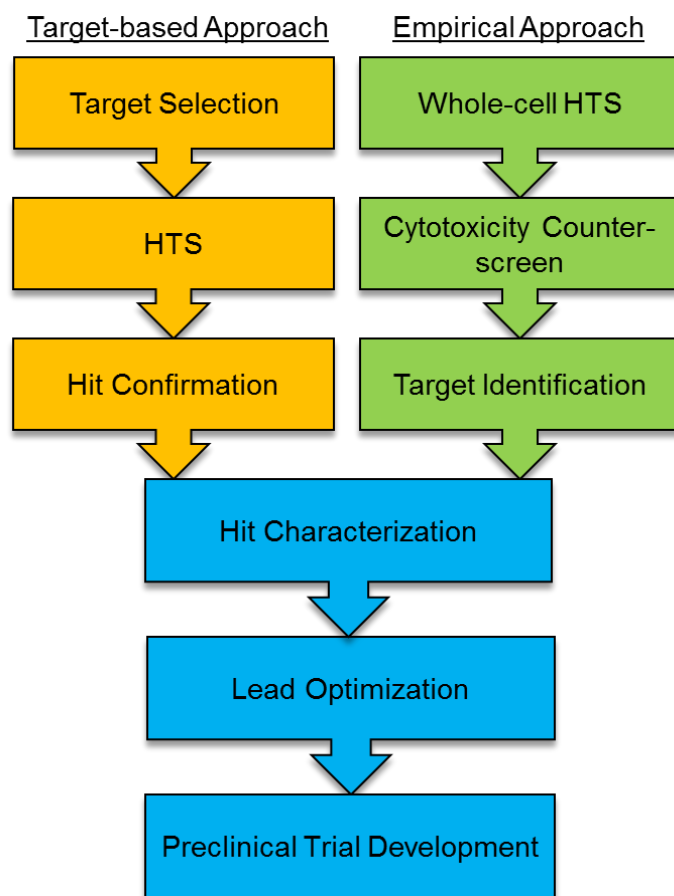


Figure 1. Workflow for two methods of antimicrobial drug discovery.

In 1995, the DNA sequence for the genome of *Haemophilus influenza* was completed [125], promising new approaches for the challenges in drug discovery and renewed excitement by various pharmaceutical companies such as GlaxoSmithKline (GSK), Merck, Pfizer, and Wyeth [126]. Before then, few genome maps were publicly available, leaving researchers with the ability to only examine small areas of a specific sequence. With increased computational power and the availability of whole genomes, bioinformatics analysis allowed for the examination of new gene targets from various species [127]. Screening strategies then shifted to target-based drug design. The initial strategy was to focus on identifying as many potential targets available within a genome and one could then work to inhibit the targets by a high-throughput screening (HTS) of drug-like compounds instead of blindly searching for compounds with antibacterial activity [97]. One prime example of the boom in research during the genomic age was that of GlaxoSmithKline (GSK). From 1995-2002, GSK attempted to identify a compound with Gram-positive or broad-spectrum antimicrobial activity. Using *Streptococcus pneumoniae* as the primary comparison genome, GSK analyzed genes across five pathogens for highly-conserved sequences resulting in more than 350 *S. pneumoniae*, *S. aureus*, and *H. influenzae* candidate target genes. From the 350 potential genes, 127 were classified as essential *in vitro* in at least one of the three organisms and 64 targets displayed attenuated growth *in vivo*. The company screened targets against more than 250,000 compounds from their library [126]. The screening led to 16 positive ‘hits’, five which went on to ‘lead’ compounds. Of the five targets (FabH, FabI, peptide deformylase, methionine tRNA synthetase, and phenylalanine tRNA synthetase) only one, peptide deformylase, led to an optimized lead molecule (GSK 1322322) that successfully went on to Phase II clinical trials [124, 126, 128]. The disappointing results

produced during these large screens were partially due to the newness of the era, the lack of knowledge of suitable targets and the feasibility of the HTS [97, 126].

Current strategies shifted towards a rational structure-based approach due to technological advances in genomic sequencing, high resolution protein structures and computational modeling. This led to a better understanding of the bacteria and target proteins [97, 129-131]. The basic concept is that the structural properties of a specific target relates to its biological activity and modification of the target will have therapeutic effects. Computer-aided molecular modeling of protein structures allows for virtual HTS screening of compounds rationally determined to interact with the target [131]. As a result, the initial trial-and-error of drug discovery is reduced as inhibitors are selected based on structural properties and there are increases in the screening efficiency since compounds have a higher chance of binding [129]. This approach was successful in the relatively rapid design of Iclaprim, a structural derivative of trimethoprim. Trimethoprim resistance in *S. aureus* is a result of a single amino acid substitution of the target dihydrofolate reductase (DHFR), resulting in a 20-fold decrease in affinity due to the loss of a hydrogen bond. Through molecular modeling of the mutated DHFR, Iclaprim was designed with a trimethochromene side chain in place of the original trimethophenyl resulting in a 20-fold increase in binding affinity, maintaining potency [97, 132].

IV. Summary

Currently, supplemental therapy for destructive periodontal disease can include the application of broad-spectrum antibiotics [133-135]. However, we now know that this is not ideal as it can alter the essential microbiome and runs the risk of promoting antibiotic resistance. While the oral cavity is a complex poly-microbial environment, it is evident that the interaction of specific

species disrupts the oral homeostasis, triggering the onset of periodontal disease. The hypothesis of *P. gingivalis* being a key player in the initiation of periodontitis suggests the potential for novel narrow-spectrum therapeutics. By targeting the key pathogens, it may be possible to treat periodontitis while allowing for the recolonization of the beneficial, healthy flora. In this study, we present a unique approach to predict essential gene targets selective for the periodontal pathogen within the oral environment. We then utilize a rational-based approach to identify small-molecule inhibitors for the potential development of antimicrobial agents against *P. gingivalis* and periodontal disease. We believe by using *P. gingivalis* and periodontal disease as a model we can begin to develop an approach for pathogen-specific antimicrobial drug discovery.

V. Outline

Following the general introduction and summary, the main goals of this dissertation are divided into three chapters. Background, methodologies, results and discussion relating to the specific aim are reported for each chapter.

Chapter One focuses on the development and verification of the framework for a novel essential gene prediction model. The benefit of the model is displayed through the quick and accurate selection of a potential species-selective drug target in the Gram-negative periodontal pathogen, *P. gingivalis*.

Chapter Two continues the evaluation of the potential drug target identified in Chapter One. The application of a CBDD method is used to identify new target specific inhibitors.

Chapter Three details the biochemical screening from the compounds identified in Chapter Two. A target-based and whole-cell based approach is used to assess *in vitro* activity.

This is concluded with an overall summary and conclusion.

Chapter One

Applying a Novel Framework for the Prediction of Essential Genes as Antimicrobial Targets

Victoria N. Stone, Brian A. Klein, Linden Hu and Ping Xu

Victoria N. Stone performed the prediction, the analysis of data and the preparation of the following manuscript. Dr. Brian A. Klein performed the mutagenesis experiments and aided in analysis of data. Dr. Ping Xu and Dr. Linden Hu acted as advisors.

Background

With the onset of the genomics era came the availability of completed genome sequences and put new focus on the study of essential genes that held the promise of a wealth of biological and genomic data. Identification of these genes enhances our knowledge of gene regulatory networks [136] and aids in our understanding of cellular functions [137]. They can provide clues to the origins of life through the examination of the minimal genome [138] and evolutionary divergence [136, 137, 139]. Additionally, the understanding of the minimal number of genes required to sustain life supports research for the developing field of synthetic biology [138, 140]. Essential genes, by definition, are critical for the growth and/or survival of an organism.

Antimicrobials target key metabolic and cellular functions, making essential genes attractive targets for drug discovery. Hu *et al.* was able to successfully identify drug targets for the fungus *Aspergillus fumigatus* from *in vivo* mouse model essential gene studies [141]. This has become more important with increasing rates of antibiotic resistance. Furthermore, the discovery of genome-wide essential gene data can present previously unexplored therapeutic targets [99, 142].

To date, a variety of approaches have been utilized to identify essential genes in a number of organisms. Large-scale experimental approaches such as systematic single gene deletions [139, 143-147], random transposon mutagenesis [148-151] and RNA interference [152, 153] utilize gene disruptions to evaluate the viability of the organism. However, these approaches can be extremely time-consuming, expensive and technically difficult with certain microbial systems. Due to the large investment, comprehensive experimental essential gene data is not readily available for the majority of sequenced genomes [154]. Currently only 31 species have genome-wide experimental data that has been deposited into the Database of Essential Genes (DEG) [155]. Computational assessment of bacterial systems offer alternatives to experimental

procedures. *In silico* methods rely on information garnered from experimentally identified essential genes to predict essentiality based on orthologous genes, protein expression, phylogenetic conservation, similar GC content and codon usage [136]. These bioinformatics approaches have recently become standard for the prediction of essential genes. Unfortunately, orthologous genes account for only a fraction of the total genome and as genetic diversity increases in phylogenetically unrelated organisms, prediction accuracy decreases due to variation in gene sequences, alternate cellular pathways and genetic redundancy [136, 156]. In addition, experimental essential gene data, which *in silico* approaches are based on, is only available in a fraction of species and differences in experimental parameters (bacteria grown in rich versus minimal medium), mutagenesis techniques (site-directed versus random transposon insertion) or simply the definition of essential has led to a lack of consensus, making predictions difficult [140, 142].

Previously, we conducted a genome-wide essential gene identification study in the Gram-positive, early oral colonizer *Streptococcus sanguinis* using systematic single and double gene deletions [139]. These essential genes were grouped by specific categories based on their KEGG functional annotation. Once linked together, we were able to create a model of essential pathways and determined that essential genes were conserved within three major categories of biological function: maintenance of the cell envelope, energy production and processing of genetic information. These three features also follow the ideal for what is considered the core of the hypothetical minimal genome encompassing the basic components necessary for cell survival [138, 140, 157]. Studies show that although essential genes themselves may diverge or change, the essential function remains relatively unaltered [140]. From these observations we established a framework to predict essential genes based on genome annotation and function. Our prediction

does not rely on sequence features among different species, but a gene's contribution to an essential end product or the overall essential function that it plays. We believe that the reliance on genomic function instead of sequence addresses inconsistencies of essential gene data among different bacterial species as a result of genetic diversity while still allowing for an accurate prediction [140]. Furthermore, as a predictive method our approach simply requires a sequenced and annotated genome, bypassing the need for previous experimental data from closely related species. When lacking comprehensive essential gene datasets, we believe our approach provides a quick and accurate prediction of essential genes within the genome of different organisms. This could be extremely beneficial, especially when concerning drug target identification of emerging pathogens or difficult to culture organisms. Knowledge based on metabolic or gene functional networks have increasingly become an invaluable tool [158]. This was evident when a previous genome-scale study with metabolic networks, similarly based on genome information, showed high accuracy in the prediction of amino acid preference and by-product secretion rates [159].

In this study, we demonstrate the framework for our essential gene predictive model is both accurate and useful for the identification of antimicrobial targets. First, using KEGG pathways and genome annotation, we identify relevant pathways and predict essential genes for the periodontal pathogen *Porphyromonas gingivalis*. The selection of *P. gingivalis* to evaluate our accuracy was valuable as it is phylogenetically distinct from *S. sanguinis*, from which the prediction model was derived. Second, we cross-validate our prediction with experimental data. During our study, *P. gingivalis* lacked a large-scale essential gene study. However, a genome-wide Mariner-based transposon mutagenesis study was recently carried out [Klein BA, unpublished], allowing us to verify the accuracy of our prediction. Third, we identify *P. gingivalis*-selective essential gene targets, through the comparison of essential gene data sets

between the primary, non-pathogenic colonizer, *S. sanguinis* [139] and *P. gingivalis*. Last, based on “druggable” target criteria, we select a promising candidate target to experimentally characterize for future drug development studies.

Materials and Methods

Bacterial Strains, plasmids and growth conditions. *Porphyromonas gingivalis* W83 strain was cultured anaerobically (10% CO₂, 10% H₂, and 80% N₂) at 37 °C in tryptic soy broth (TSB) (Becton Dickinson, Franklin Lakes, NJ) supplemented with 1 µg/ml menadione and 5 µg/ml hemin. When appropriate, erythromycin (Fisher Scientific, Fair Lawn, NJ) was used at a concentration of 5 µg/ml. Plasmid pVA2198 (Richmond, VA) [160] was used to isolate the erythromycin resistance cassette.

Prediction of *P. gingivalis* essential genes. The prediction of essential genes was based on data derived from Xu *et al.* [139]. Using the Kyoto Encyclopedia of Genes and Genomes (KEGG) database [161] for the genome-wide analysis of *P. gingivalis* strain W83, pathways such as glycolysis and fatty acid biosynthesis that were related to the three broader categories previously described (maintenance of the cell envelope, energy production and processing of genetic information) were manually examined and a network of predicted essential pathways was determined. We identified genes whose function singularly contributed to the formation of an essential end product for each of the pathways within our described network. Alternative pathways would be noted for any essential pathways within our network and paralogs/isozymes of predicted essential genes would be noted and listed as putatively essential. The assigned KO numbers and path numbers of genes predicted to be essential were listed and grouped by biological function. Multiple KO numbers or path numbers from KEGG were often assigned for

a single gene if it was involved in different pathways. Predicted essential gene functions wrongly annotated or missing annotation would be predicted by BLAST comparison of genes with similar gene functions from phylogenetically similar species.

Gene functions assignment. To determine putative gene function, protein sequences were searched against annotated *Porphyromonas gingivalis* strain W83. The Cluster of Orthologous Groups (COG) annotation [162] was downloaded from the NCBI database and essential pathways were analyzed via the KEGG database [161]. Putative gene functions and COGs were recorded for predicted essential genes and the total number of genes in the genome. Ratios of the number of essential genes for each COG and the total number of genes belonging to a specific COG category were calculated.

Comparison of Essential Gene Homologs. To identify essential genes conserved in *P. gingivalis* W83, a comparative genomics analysis was performed. Thirty-one other organisms with essential genes annotated in the DEG v10.0 [163] were obtained and protein sequence conservation was determined by BLASTP. Significant matches ($E < 1e-5$) were analyzed to find homologs to *P. gingivalis* genes. Protein sequence conservation was calculated by percent amino acid identity.

Multiple sequence alignment. For the prediction of the substrate binding site, the protein knowledge database, UniProtKB/Swiss-Prot (www.uniprot.org) [164], was referenced for organisms with completed enzymatic and functional data for *m*-Ddh. This included *Corynebacterium glutamicum* ATCC 27405, *Lysinibacillus sphaericus*, *Bacteroides fragilis* ATCC 25285, *Clostridium thermocellum* ATCC 13032 and *Ureibacillus thermosphaericus* (Uniprot ID: P04964, Q9KWR0, Q5L9Q6, A3DDX7, G1UII1). Complete protein sequences

were obtained from the National Center for Biotechnology Information (NCBI, www.ncbi.nlm.nih.gov/) database. Sequence alignment analyses were then performed using the ClustalW multiple alignment [165] via the BioEdit program v7.1.7. The alignment figure was generated in Esript 3.0 [166] for visualization.

Primer design and recombinant PCR product construction. Primer design was based on the method of Xu *et al.* [139]. Two sets of primers, F1/F2 (5'- CTC CGA ATA GCA AAC ATC TAC TG -3' and 5'- GAA AAA TTT CAT CCT TCG TAG TCG AGC AGC CAT GCG C -3') and F3/R3 (5'- GGG CAA TTT CTT TTT TGT CAT TTG TCA AAT CTG GGG G -3' and 5'- GAT AAT CAT GCT TCG GAG ATG -3'), were designed to amplify a 1.2kb sequence upstream and downstream, respectively, of the *P. gingivalis* target gene. A third primer set, F2/R2 (5'- GCG CAT GGC TGC TCG ACT ACG AAG GAT GAA ATT TTT C -3' and 5'- CCC CCA GAT TTG ACA AAT GAC AAA AAA GAA ATT GCC C -3') was designed to amplify the 800 bp erythromycin resistance cassette (Erm^R) containing the *ermF* gene from plasmid pVA2198. To minimize polar effects from mutagenesis, primers were designed to include stop codons within frame and the antibiotic resistance cassette was designed to run in the same orientation as the target gene to ensure transcription. The three PCR fragments were created using F1/R1, F2/R2 and F3/R3. All PCR reactions were performed with an initial denature of 98 °C for 30 sec, 30 cycles of 98 °C for 10 sec, 56 °C for 30 sec, 72 °C for 36 sec and a final extension of 72 °C for 7 min. The PCR products were purified using QIAquick PCR Purification Kit (Qiagen, Valencia, CA); the three fragments were then combined in equal amounts and amplified again using F1 and R3 primers to generate the final linear recombinant product. The PCR reaction was performed with an initial denature of 98 °C for 30 sec, 30 cycles of 98 °C for 10 sec, 56 °C for 30 sec, 72 °C for 1 min 36 sec and a final extension of 72 °C for 7

min. Phusion High-Fidelity Taq DNA polymerase (New England Biolabs, Ipswich, MA) was used in all reactions.

Transformation. The electroporation method was adapted from Fletcher *et al.* [167]. Briefly, 0.2 ml of an actively growing culture of *P. gingivalis* was used to inoculate 2 ml of TSB supplemented with yeast extract, hemin and menadione, which was then incubated overnight at 37 °C. Five ml of the same medium pre-warmed to 37 °C was then inoculated with 0.5 ml of the overnight culture and was incubated for an additional 4 h ($OD_{600} \approx 0.7$). The cells were harvested by centrifugation at $6,000 \times g$ for 15 min at 4 °C and washed twice in 10 ml of ice-cold electroporation buffer (10% glycerol, 1 mM $MgCl_2$). The final cell pellet was re-suspended in 0.5 ml of electroporation buffer. A 100 μ l sample of re-suspended cells and 5 μ g of DNA were placed in a sterile electrode cuvette (0.2-cm gap). The cells were then pulsed with a Bio-Rad (Hercules, Calif.) gene pulser at 2,500 V for 9.5 ms and incubated on ice for 5 min. The cell suspension was then added to 0.6 ml of TSB broth supplemented with yeast extract, hemin and menadione and incubated for approximately 16 h. A 100 μ l sample was plated on 5% sheep blood agar medium containing erythromycin and incubated anaerobically at 37 °C for 5 to 10 days.

Results

Defining our framework of essential gene prediction. The definition of gene essentiality can vary between researchers, by experimental parameters and/or growth conditions. Genes that are essential for growth in minimal medium or during *in vivo* conditions may not necessarily reflect what is essential for survival in ‘ideal’ laboratory conditions. Also, a deletion or mutation of a defined essential gene may not be lethal but may lead to decreased fitness [168]. These

discrepancies are clear in the model organism *Escherichia coli*. Studies performed by several laboratories, showed clear differences in their definition and number of experimentally identified essential genes. Gerdes *et al.* performed genome-wide transposon mutagenesis in 2003, where they identified 609 essential genes [169], the Keio collection is currently defined as 296 essential genes [143] and a metabolic study described 119 genes as being conditionally essential [170]. In this study we focused our prediction on genes we considered to be necessary for growth in rich media, implying that certain vitamins, amino acids and nucleotide precursors would be provided. These genes represent the ‘core’ of the cell’s biological and cellular functions and would be considered essential in typical experimental *in vitro* conditions [171]. Furthermore, as these core essential functions are crucial to bacterial fitness, they represent a set of genes less affected by evolutionary processes [172-174].

From experimental studies in *S. sanguinis* SK36 [139], we saw that, while essential genes were distributed in various pathways, they could be associated through key biochemical pathways or functions. Genes required for peptidoglycan biosynthesis, terpenoid backbone biosynthesis, glycerophospholipid metabolism and fatty acid biosynthesis all related to the maintenance of the cell envelope. The pentose phosphate pathway and glycolysis provided energy. Nucleotide biosynthesis, DNA replication, cell division proteins, and protein biosynthesis were associated with the processing of the cell’s genetic information. In order to predict essential genes, we identified essential metabolic pathways corresponding to the three core biological functions. The functional genome annotation was obtained from the KEGG database and assigned to genes within the pathways. Our approach defined essential as a gene whose functional role individually contributed to the formation of a crucial chemical compound. Therefore the single deletion or mutation of the gene would result in a non-viable cell.

Prediction of essential genes by gene annotation in *P. gingivalis* strain W83. To predict essential genes in *P. gingivalis* strain W83, we manually examined our essential biochemical pathway maps through the KEGG database, a workflow schematic is shown in Fig. 2. Annotated genes were viewed for each pathway and from 1909 protein coding genes, 212 were predicted to be essential (Fig. 3, for complete list, see Appendix, Supp. Table 1). During our prediction, we noted several instances of paralogs or potential isozymes (PG1852/PG0223, exonuclease, DNA polymerase III subunit epsilon; PG0121/PG1258, DNA binding protein HU; and PG1940/PG0933, elongation factor G). Many paralogous genes code for proteins that have redundant functions, which is why a majority of computational models predict them to be non-essential. However, it has also been noted that paralogous genes can have distinct roles through functional divergence [175]. For example, *N. meningitidis* possess paralogous genes for glyceraldehyde-3P dehydrogenase (NMB2159 and NMB0207). However, only NMB2159 was determined to be essential, as NMB2159 is involved in glycolysis and NMB0207 is involved in gluconeogenesis [176]. As experimental data would be needed to determine if both or one gene was essential, we listed paralogs as potentially essential. Hypothetical proteins were left out of our analysis unless a conserved motif and putative function was noted in the KEGG database. One such example was PG2046, which was annotated as a hypothetical protein in the KEGG database. Based on our database searches the gene contained a conserved tRNA methyltransferase motif and had orthologous matches to other tRNA(Ile)-lysine synthases. This indicated that the gene might have a putative essential function in tRNA processing.

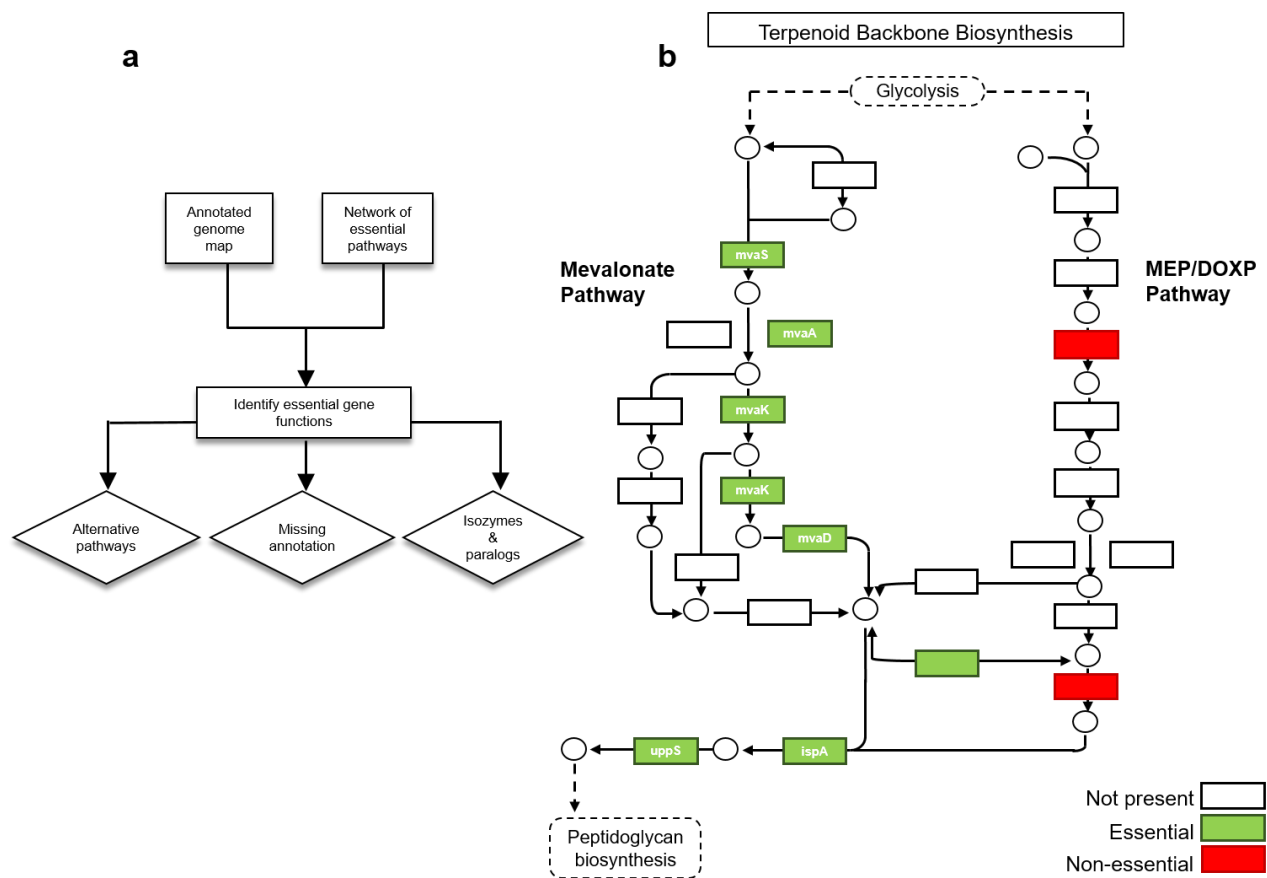


Figure 2. Workflow schematic for the prediction model framework.

(a) A strategy to predict essential genes. **(b)** Example of an essential pathway from *S. sanguinis* strain SK36. Circles represent key intermediates; green boxes represent essential genes; red boxes represent non-essential genes; white boxes represent genes not present within the genome of *S. sanguinis*.

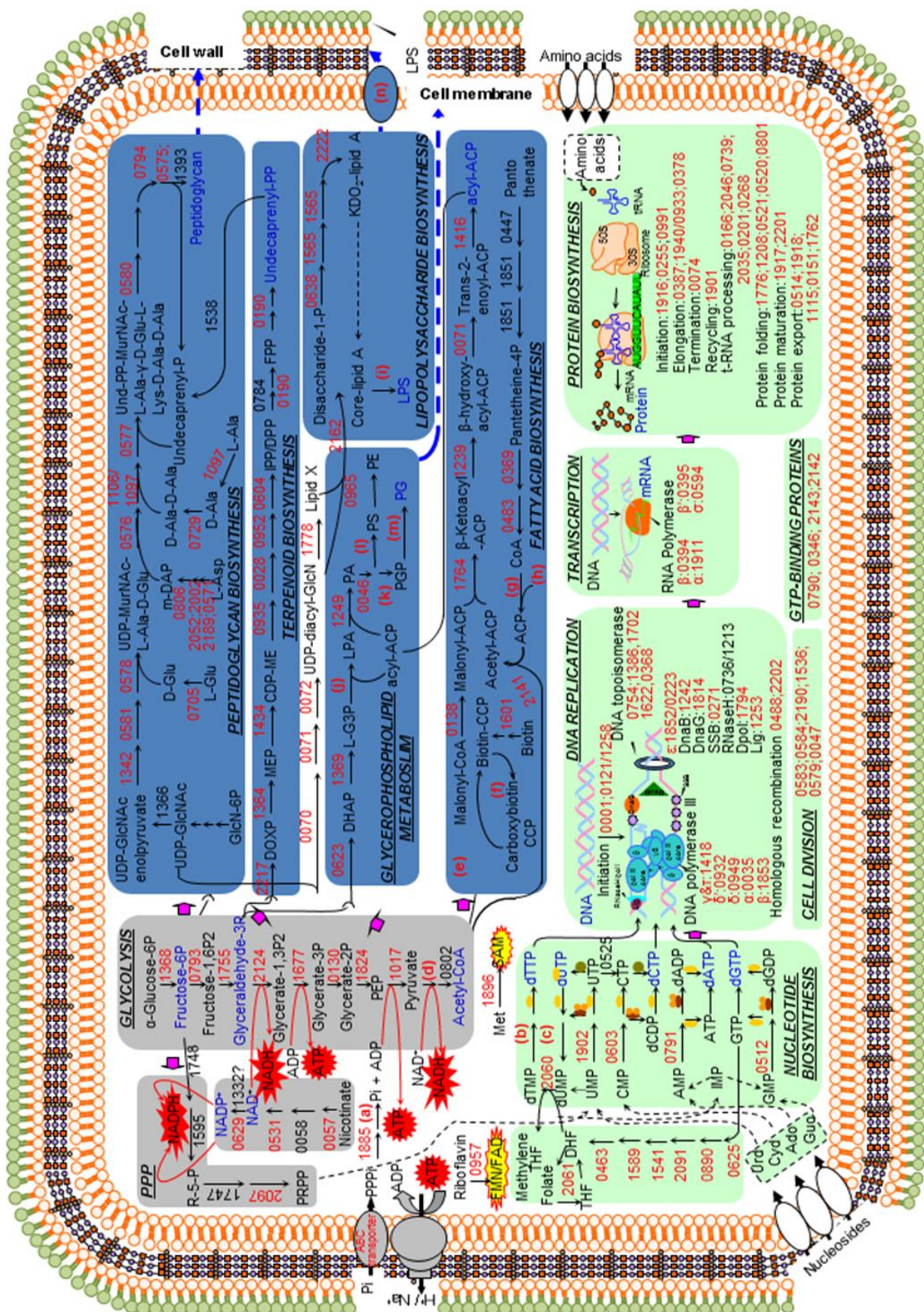


Figure 3. Predicted essential genes and corresponding pathways for *P. gingivalis* strain W83.

The three biological categories associated with essential genes, previously determined from experimental studies in *S. sanguinis* SK36, are indicated by color: dark blue, maintenance of cell envelope; grey, energy production; light green, processing of genetic information. Predicted essential gene numbers (PG#) are shown in red, predicted non-essential genes are shown in black. Solid arrows indicate enzymatic reaction, dashed arrows dashed arrow, multistep pathway; blue dashed arrow, products involved in cell wall and membrane formation; block arrow, product from one pathway serving as input to another pathway; oval with bold arrow, transporter; slash, separating paralogs. Ado, adenosine; Cyd, cytidine; DAG, 1,2-diacylglycerol; DHAP, dihydroxyacetone phosphate; DHF, dihydrofolate; L-G3P, sn-glycerol 3-phosphate; Guo, guanosine; LPA, lysophosphatidic acid; LTA, lipoteichoic acid; PA, phosphatidic acid; PS, phosphatidylserine; PE, phosphatidylethanolamine; PGP, phosphatidylglycerophosphate; PG, phosphatidylglycerol; PPP, pentose phosphate pathway; LPS, lipopolysaccharide; THF, tetrahydrofolate; Urd, uridine. Genes assumed to be essential based on prediction, but not annotated in the KEGG database, are represented by red letters (see Table 1). Predicted essential genes associated with ribosome and aminoacyl-tRNA biosynthesis are not indicated here. For the complete essential gene prediction see Appendix, *Supp. Table 1*.

From our assessment of the core essential pathways, we knew certain genes should be critical to the formation of certain compounds. However, during our analyses we saw that a small fraction of genes were missing annotations for certain essential functions in the KEGG database, noted in Fig. 3 as lowercase letters. For example, in the glycolysis pathway, the gene responsible for the conversion of pyruvate to acetyl-CoA is not noted. To identify these missing genes, we compared the genome of *P. gingivalis* against annotated genes in the NCBI database using a BLASTP analysis. Proteins with e-values of 1×10^{-8} or less were then prioritized and missing genes were identified by their potential function or motif (Table 1, Fig. 3).

There is no direct correlation between the number of essential genes and the total number of genes in the genome as different experimental approaches or conditions may derive various outcomes. However, on average, essential genes account for around 10-20% of the genome. *S. sanguinis* SK36 with 2270 protein encoding genes has 218 essential genes or 9.6% of the genome [139], *S. pneumoniae* D39 with 2,046 ORFs has 244 essential genes or 11.9% [146, 147] the 4,291 protein encoding genes in *E. coli* MG1655, 620 were assessed to be essential or 14.4% [169]. *B. subtilis* 168 resulted in 271 essential genes from a total ORF of 4,099 genes, accounting for 6.8% of the genome [145] and *P. aeruginosa* contains 678 putatively essential genes from 5,500 genes or approximately 12% [177]. Our prediction of 212 essential genes accounted for 11% of the genome, fitting with the expected percentage of essential genes within a microbial genome. Nonetheless, this prediction may not be comprehensive. Some genes could have been overlooked due to a lack of biological data annotated in the KEGG database, for instance the lack of annotation for hypothetical genes.

Table 1. Non-annotated predicted essential genes in *P. gingivalis* strain W83.

Figure letter	Gene function	Probable Gene	KEGG ID	E-value
a	pyrophosphatase	type I phosphodiesterase/nucleotide pyrophosphatase	PG0126	0
b	thymidylate kinase	thymidylate synthase	PG2060	6.00E-105
c	thymidylate kinase	thymidylate synthase	PG2060	6.00E-105
d	pyruvate dehydrogenase	alpha keto acid dehydrogenase complex, E3 component, lipoamide dehydrogenase	PG0802	0
e	acetyl-CoA carboxylase	methyl-CoA decarboxylase subunit alpha	PG1612	0
e	acetyl-CoA carboxylase	methyl-CoA decarboxylase subunit alpha	PG1612	0
e	acetyl-CoA carboxylase biotin carboxyl carrier protein	oxaloacetate decarboxylase	PG0249	0
f	acetyl-CoA carboxylase biotin carboxylase	oxaloacetate decarboxylase	PG0249	0
g	4-phosphopantetheinyl transferase	4-phosphopantetheinyl transferase	PG0273	4.00E-116
h	acyl carrier protein	UDP-3-O-[3-hydroxy-myristoyl] glucosamine N-acyltransferase	PG0072	0
i	O-antigen ligase	hypothetical protein	PG0109	0
j	glycerol-3-phosphate acyltransferase	hypothetical protein	PG0728	6.00E-100
		hypothetical protein	PG1051	0
k	CDP-diacylglycerol-inositol 3-phosphatidyltransferase	hypothetical protein	PG2070	1.00E-05
l	phosphatidylserine synthase	hypothetical protein	PG1104	5.00E-38
m	phosphatidylglycerophosphatase A	hypothetical protein	PG1738	0
n	lipopolysaccharide export system protein	ABC transporter ATP-binding protein	PG0628	6.00E-104

Functional analysis and examination of predicted essential genes by biological categories.

To ensure that our essential genes followed our model, we next examined them to determine if they were associated with the three major biological categories. The predicted essential genes in *P. gingivalis* were grouped into a pathway based on the KEGG pathway orthology (KO number) and the final end product. Of the 212 predicted genes, 46 genes fell into the maintenance of the cell envelope category. About 32% (7% of total predicted essential genes) of those genes were directly related to peptidoglycan biosynthesis while about 19% (4% of total) were related to fatty acid biosynthesis. Sixteen genes were involved in energy production with about 69% (5% of total) related to glycolysis. The majority of the predicted essential genes (148 genes) were grouped into the processing of genetic information category. About 9% (6% of total) were related to nucleotide biosynthesis including both and pyrimidine metabolism and 33% (24% of total) were related to ribosomal biosynthesis. Two genes were involved in the synthesis of cofactors: PG0957, a riboflavin biosynthesis protein involved in FMN/FAD biosynthesis, and PG1896, annotated as S-adenosylmethionine (SAM) synthetase. Both of these genes fit into a variety of metabolic roles, producing energy which feed into other pathways (Fig. 4a).

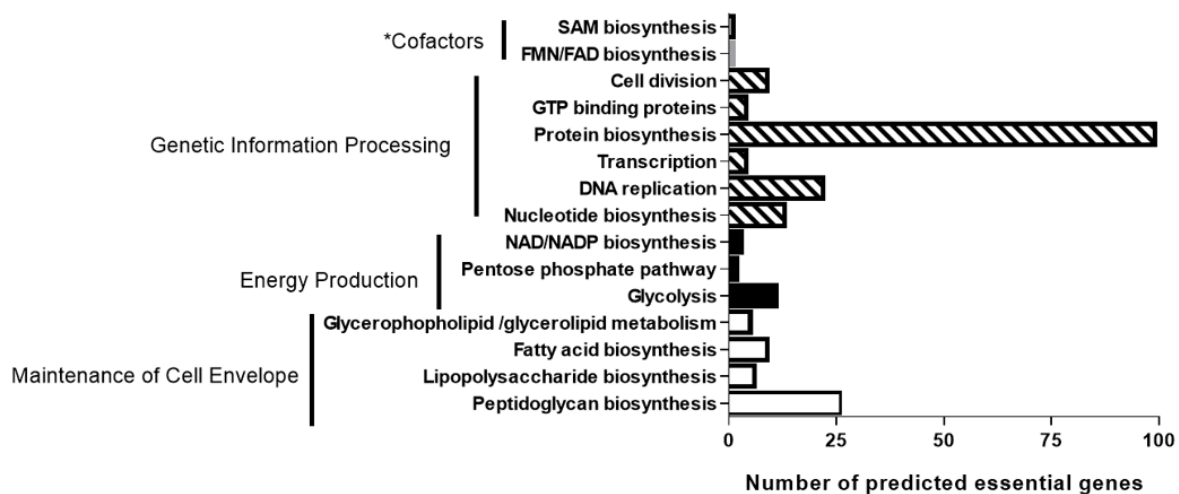
To further study the functional characteristics of our predicted essential gene list, Clusters of Orthologous Groups of proteins (COG) [162] classifications were examined. COGs are defined through protein sequence comparison of completed genomes and functional categories such as RNA processing and modification, cell motility and lipid metabolism are grouped. A high number of genes predicted to be essential, were involved in translation, ribosomal structure and biogenesis (group J; Fig 4b) following other essential gene and minimal genome studies [143, 157, 162]. These COG functions belong to the information storage and processing category which includes critical cellular functions that require multiple proteins such as 30S and 50S

ribosomal proteins, cell division proteins and translation factors. However, the larger total proportion of genes belonged to COG functional groups that are related to metabolism (group C, E, F, G, H, I and Q; Fig 4b). COG assignments are influenced by genome annotation. Our prediction was mostly lacking hypothetical proteins without a functional annotation; therefore, only eight genes did not have a clearly defined COG function and are classified by general function. This included GTP-binding proteins *era*, *obgE*, *engB* and *engA*, acyl carrier proteins involved in fatty acid biosynthesis *fabG* and *fabK*, an oxidoreductase involved in peptidoglycan biosynthesis, and the ribonuclease *Z eluC*. This COG distribution is consistent with genes necessary for *in vitro* growth under laboratory conditions. Alteration in the environmental conditions would lead to changes in genes required for survival. Studies show that cells grown under chemically defined minimal media or during *in vivo* conditions were not enriched for translation, ribosomal structure and biogenesis (group J), but for amino acid transport/metabolism (group E), nucleotide transport/metabolism (group F) and energy production/conversion (group C) [178]. This suggests there are different needs for different environmental conditions.

Cross-validation of prediction with experimental validation of essential genes in *P.*

***gingivalis* strain W83.** To evaluate our model, we performed a comparison of the predicted essential with putative experimental essential gene data. A transposon mutant library was generated in *P. gingivalis* strain W83 using a pSAM_Bt-based Mariner mutagenesis vector as previously carried out in Klein et al. for the genome-wide mutagenesis of *P. gingivalis* strain ATCC 33277 [149]. Approximately 20,000 colonies were pooled following multiple separate transpositions. Strain W83 proved difficult to obtain transposon mutants, which may be a result

a



b

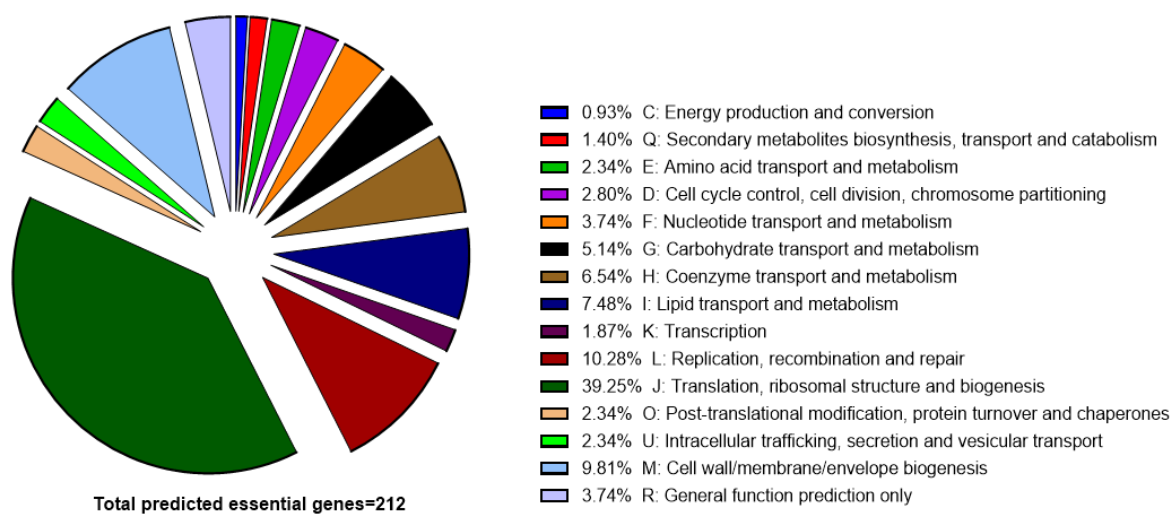


Figure 4. Pathway association and COG distribution for predicted essential genes in *P. gingivalis* strain W83.

- (a) Distribution of predicted essential genes divided by pathway and three conserved biological functions.
- (b) The number of predicted essential genes and the percent distribution for the total number of predicted essential genes were calculated for the individual COG function.

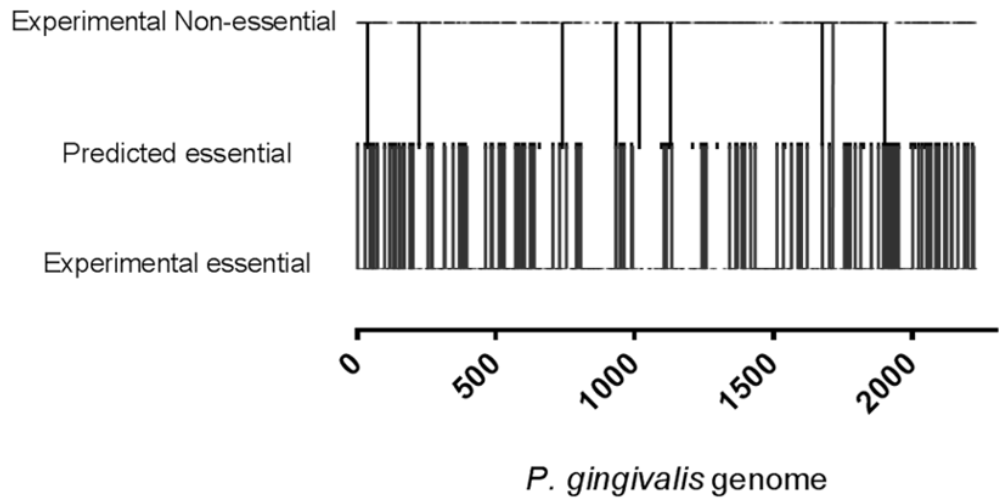
of known capsule and fimbriae structures, characteristic of this strain. A small degree of resistance to the selective antibiotic (Erm) was also observed, which may point to survival of a mutant clone being correlated with the expression level of the Erm cassette. Following Illumina sequencing of two technical replicates, it was determined that about 12,000 distinct insertion mutants were present in the strain W83-background mutant pool. Around 12,000 insertions is 5-fold coverage of the genome, which is sub-optimal for characterization of gene essentiality. Based on the limited number of insertions relative to the 36,000 distinct insertions in the ATCC 33277-background, we decided to focus on grouping genes as either containing zero insertions or greater than two insertions, corresponding to putatively essential and putatively non-essential respectively (Appendix, Supp. Fig. 1) With this grouping guideline, 759 putatively essential and 680 putatively non-essential genes were noted. It is important to note that the guidelines for determining putatively essential and non-essential in strain W83 were not as stringent as those in the previous study for strain ATCC 33277 due to the different library complexities.

From 212 predicted essential genes, 173 matched the experimental data, giving us a prediction accuracy of 81.6% (Fig 5, Supp. Table 1). Of the predicted essential genes not experimentally validated, three of these genes (PG0223, DNA polymerase III subunit epsilon; PG0121, DNA binding protein HU; PG0933, elongation factor G) may have been essential if not for possible isozymes or paralogs to predicted essential genes that were validated (PG1852, PG1258, PG1940, respectively). A large subset was annotated as hypothetical and therefore could not be predicted by our model without an assigned cellular function. A majority of these genes coded for functions that play multiple, non-specific roles. For example, there were several ABC transporters, ATP-binding proteins, glycosyltransferase and putative transcriptional regulators. In addition, many of these genes were not linked or associated with biochemical

pathways in KEGG. Lacking this information made it difficult to assess the essential role the gene would play to the overall cell survival. KEGG annotations are not necessarily the most up-to-date.

Although we were interested in the prediction of essential genes, to further validate the accuracy of our model, we compared the predicted essential genes to the experimentally determined non-essential genes to see if there was any overlap between the two data sets. From 680 observed non-essential genes in strain W83 (Appendix, Supp. Table 1), only 7 genes (or 3% total) were predicted to be essential by our model (Fig. 5). These few false positives were again due to the presence of potential isozymes, paralogs or alternative pathways. PG0223 and PG1852, for example, both encode exonuclease DNA polymerase III subunit epsilon involved in DNA replication. While PG1852 is putatively essential, PG0223 is not. Mis-predictions may also come from lack of, or incorrect, annotation. In glycolysis, the conversion of PEP to pyruvate by the irreversible enzyme pyruvate kinase should be essential for energy production by the cell. However, the annotation for pyruvate kinase was missing in KEGG for *P. gingivalis*. A reversible enzyme with a similar function, pyruvate phosphate kinase, is found in some bacteria for the conversion of pyruvate to PEP and this enzyme was present and annotated as PG1017. Due to the lack of annotation for pyruvate kinase, we predicted pyruvate phosphate kinase to play that essential function. As PG1017 was experimentally shown to be non-essential, it is likely pyruvate kinase has not been noted within the database or another enzyme must be present. A small number of genes (32 out of 212) were unaccounted for in both the non-essential and essential experimental data. As mentioned earlier, due to complexities in the experimental conditions, some genes could not be accurately assigned to a category and therefore left out of the analysis.

a



b

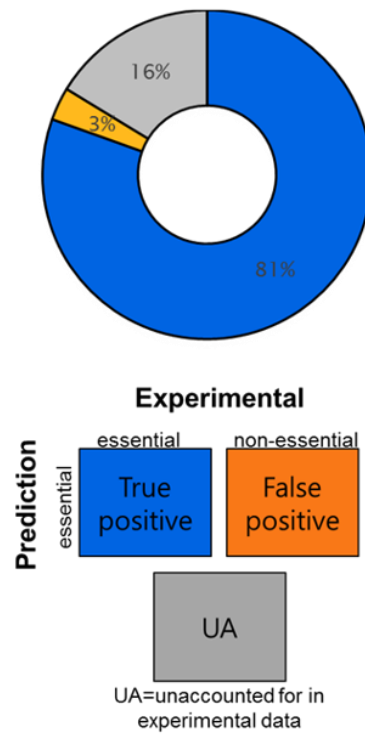


Figure 5. Cross-validation of prediction to experimental essential gene data.

(a) The number of essential genes shared between those predicted and the experimentally identified essential and non-essential gene data. **(b)** The comparison of the number of essential genes between the predicted essential genes, experimentally identified essential genes and experimentally determined non-essential genes for strain W83 is shown. The blue circle represents the number of predicted genes correlated with experimental data, orange the number of predicted essential genes correlated with non-essential from experimental data and gray the number of predicted essential genes not accounted for in experimental data. Approximately 81% of genes predicted to be essential (173 of 212) matched the experimental essential data set, and only a small fraction (7 of 212) of genes predicted to be essential were shown to be non-essential.

Comparison of *P. gingivalis* and *S. sanguinis* essential genes. We next compared the essential gene data predicted for *P. gingivalis* strain W83 to the experimental data of *S. sanguinis* strain SK36. As the healthy oral cavity is composed of roughly 80% streptococcus species [179], we theorized that selecting orthologous essential genes in *P. gingivalis* that were absent or not essential within *S. sanguinis* would present us with potential antimicrobial targets selective within the oral cavity for periodontal pathogens. We identified 68 essential genes that were selective for *P. gingivalis* (Table 2). It was clear during our analysis that differences were mostly a result of alternative pathways or genes apart of complexes that we predicted to be essential to the overall function. A clear example is in regards to cell wall composition. Gram-negative bacteria possess an outer membrane composed of lipopolysaccharide and lipoproteins with a high lipid content. Gram-positive bacteria express lipoteichoic acid on the cell membrane and is composed of a high peptidoglycan content. This gave us a different set of essential genes for cell envelope maintenance. Another key difference lay in terpenoid biosynthesis. *S. sanguinis* uses the mevalonic acid pathway, producing terpenoids via the HMG-CoA reductase pathway. However, *P. gingivalis* uses the alternate MEP/DOXP or non-mevalonate pathway.

Identification and assessment of *meso*-diaminopimelate as a target. To assess the value of our model in drug discovery, we wanted to examine the list of *P. gingivalis*-selective essential genes for future therapeutic targets. Sequences for each gene was first searched against the human genome to eliminate genes with significant human homology. No genes within our list had a homologous sequence compared to the human genome. Based on what we know of bacterial metabolism, the lysine biosynthesis pathway presented an attractive option for pathogen-selective targeting. Lysine is a required amino acid for bacteria and especially

Table 2. Predicted *P. gingivalis*-selective essential genes.

Functional category	KEGG Pathway	KEGG KO#	GeneID	Name	Protein ID
Cell envelope	Peptidoglycan biosynthesis	K03340	PG0806	-	Gfo/Idh/MocA family oxidoreductase
		K01921	PG0729	Ddl	D-alanyl-alanine synthetase A
		K01929	PG1106	murF	D-Ala-D-Ala adding enzyme
		K01000	PG0577	mraY	phospho-N-acetylmuramoyl-pentapeptide-transferase
		K02563	PG0580	murG	N-acetylglucosaminyl transferase
		K05366	PG0794	pbp1a	penicillin-binding protein 1A
		K03587	PG0575	-	penicillin-binding protein 2
		K00215	PG2002	dapB	dihydrodipicolinate reductase
		K00928	PG2189	lysC	aspartate kinase
		K00133	PG0571	asd	aspartate-semialdehyde dehydrogenase
		K01714	PG2052	dapA	dihydrodipicolinate synthase
	Terpenoid backbone biosynthesis (MEP/DOXP pathway)	K01662	PG2217	dxs	1-deoxy-D-xylulose-5-phosphate synthase
		K00099	PG1364	dxr	1-deoxy-D-xylulose 5-phosphate reductoisomerase
		K00991	PG1434	ispD	2-C-methyl-D-erythritol 4-phosphate cytidyltransferase
		K00919	PG0935	ispE	4-diphosphocytidyl-2-C-methyl-D-erythritol kinase
		K01770	PG0028	ispF	2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase
		K03526	PG0952	ispG	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase
		K03527	PG0604	ispH	4-hydroxy-3-methylbut-2-enyl diphosphate reductase
	Terpenoid backbone biosynthesis	K00806	PG0190	uppS	undecaprenyl pyrophosphate synthetase
	Polysaccharide transporter	-	PG0117	-	polysaccharide transport protein
	Lipopolysaccharide biosynthesis	K00677	PG0070	lpxA	UDP-N-acetylglucosamine acyltransferase

		K02372	PG0071	lpxC	bifunctional UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosamine deacetylase/(3R)-hydroxymyristoyl-ACP dehydratase
		K02536	PG0072	lpxD	UDP-3-O-[3-hydroxymyristoyl] glucosamine N-acyltransferase
		K00912	PG0638	lpxK	tetraacyldisaccharide 4'-kinase
		K02527	PG1565	-	3-deoxy-D-manno-octulosonic-acid transferase
		K02517	PG2222	-	acyltransferase
	Glycerophospholipid metabolism/glycerolipid metabolism	K00057	PG1369	gpsA	glycerol-3-phosphate dehydrogenase
		K00980	PG2068	tagD	glycerol-3-phosphate cytidylyltransferase
Energy production	Glycolysis / Gluconeogenesis	K01835	PG2010	pgm	phosphomannomutase
		K04041	PG0793	fbp	fructose-1,6-bisphosphatase
		K15634	PG1513	-	phosphoribosyltransferase/phosphoglycerate mutase
		K01006	PG1017	ppdk	pyruvate phosphate dikinase
		K01610	PG1676	pckA	phosphoenolpyruvate carboxykinase
	Pentose phosphate pathway	K01619	PG1996	deoC	deoxyribose-phosphate aldolase
	Nicotinate and nicotinamide metabolism	K00763	PG0057	pncB	nicotinate phosphoribosyltransferase
		K01950	PG0531	nadE	NAD synthetase
Genetic information processing	Pyrimidine metabolism	K00384	PG1134	trxB	thioredoxin reductase
		K00945	PG0603	cmk	cytidylate kinase
	Purine metabolism/Pyrimidine metabolism	K00525	PG1129	nrd	ribonucleotide reductase
	Folate biosynthesis	K01633	PG2091	folB	dihydroneopterin aldolase
	DNA replication (DNA polymerase III)	K02342	PG1852	-	exonuclease (DNA polymerase III subunit epsilon)
		K02342	PG0223	-	exonuclease (DNA polymerase III subunit epsilon)
	Ribosome	K02919	PG1915	rpmJ	50S ribosomal protein L36

		K02916	PG0990	rpmI	50S ribosomal protein L35
		K02895	PG1927	rplX	50S ribosomal protein L24
		K02897	PG0167	rplY	50S ribosomal protein L25
		K02876	PG1919	rplO	50S ribosomal protein L15
		K02907	PG1920	rpmD	50S ribosomal protein L30
		K02888	PG0314	rplU	50S ribosomal protein L21
		K02904	PG1930	rpmC	50S ribosomal protein L29
		K02952	PG1914	rpsM	30S ribosomal protein S13
		K02956	PG1758	rpsO	30S ribosomal protein S15
		K02996	PG0376	rpsI	30S ribosomal protein S9
	Aminoacyl-tRNA biosynthesis (tRNA synthetase)	K01886	PG1951	glnS	glutaminyl-tRNA synthetase
		K01893	PG1121	asnC	asparaginyl-tRNA synthetase
		K01876	PG0153	aspS	aspartyl-tRNA synthetase
	(tRNA processing)	K04075	PG2046	-	hypothetical protein: tRNA(Ile)-lysine synthase
		K00566	PG0268	mnmA	tRNA-specific 2-thiouridylase MnmA
	Translation factors (Elongation factors)	K02519	PG0255	infB	translation initiation factor IF-2
	Protein export	K12257	PG1762	secDF	bifunctional preprotein translocase subunit SecD/SecF
	Chaperones and folding catalysts (Protein folding)	K00970	PG0801	-	poly (A) polymerase
	GTP-binding proteins	K03595	PG2142	era	GTP-binding protein Era
	Chromosome partitioning proteins (cell division)	K03590	PG0583	ftsA	cell division protein FtsA
		K03798	PG0047	ftsH	cell division protein FtsH
		K03589	PG0582	ftsQ	cell division protein FtsQ
		K03569	PG1396	mreB	rod shape determining protein MreB
		K03570	PG1395	mreC	rod shape determining protein MreC
Cofactors	Riboflavin metabolism	K11753	PG0957	ribF	riboflavin biosynthesis protein RibF

interesting is the dual role the pathway plays in lysine and peptidoglycan biosynthesis, making evident its potential for antimicrobial therapy. However, what makes lysine biosynthesis suitable for species-selective targeting is the presence of pathway variants (Fig. 6). The pathway is composed of four different branches, differing by the substrate intermediates at the branch point of L-2, 3, 4, 5-tetrahydrodipicolinate's (THDP) conversion to *meso*-diaminopimelate (*m*-DAP). The succinylase branch utilizes succinyl-CoA to generate succinylated intermediates; similarly, the acetylase branch utilizes acetyl-CoA to produce acetylated intermediates. These two variants are used by the majority of Gram-negative and Gram-positive bacteria. The aminotransferase branch, used by plants and methanococci, involves a single step amine transfer to produce the precursor of *m*-DAP, LL-DAP [180]. However, for *P. gingivalis*, *m*-DAP [181-185] is directly produced by *meso*-diaminopimelate dehydrogenase (*m*-Ddh; PG0806; GenBank ID: AAQ65966.1) in a single step [181-185]. This led us to focus on *m*-Ddh as a potential target to pursue. To further narrow selectivity, microbes from the Human Oral Microbiome Database (HOMD) with an annotated *m*-Ddh were collected and aligned against *m*-Ddh in *P. gingivalis*. Out of 315 sequenced genomes, 69 contained an ortholog to *m*-Ddh. Interestingly, these 69 species mostly included known pathogens such as *Prevotella* sp., *Tannerella* sp. and *Veillonella* species. This indicated the target could be useful across multiple pathogens contributing to periodontal diseases.

An important component in antimicrobial therapy is the ability to target critical biological processes required for bacterial cell survival. Historically, these targets have focused on key biological functions such as DNA replication, protein translation and cell wall biosynthesis [186]. While *m*-Ddh is involved in protein and cell wall biosynthesis and we predicted the gene to be essential, prior to the beginning of our study we did not know whether this was

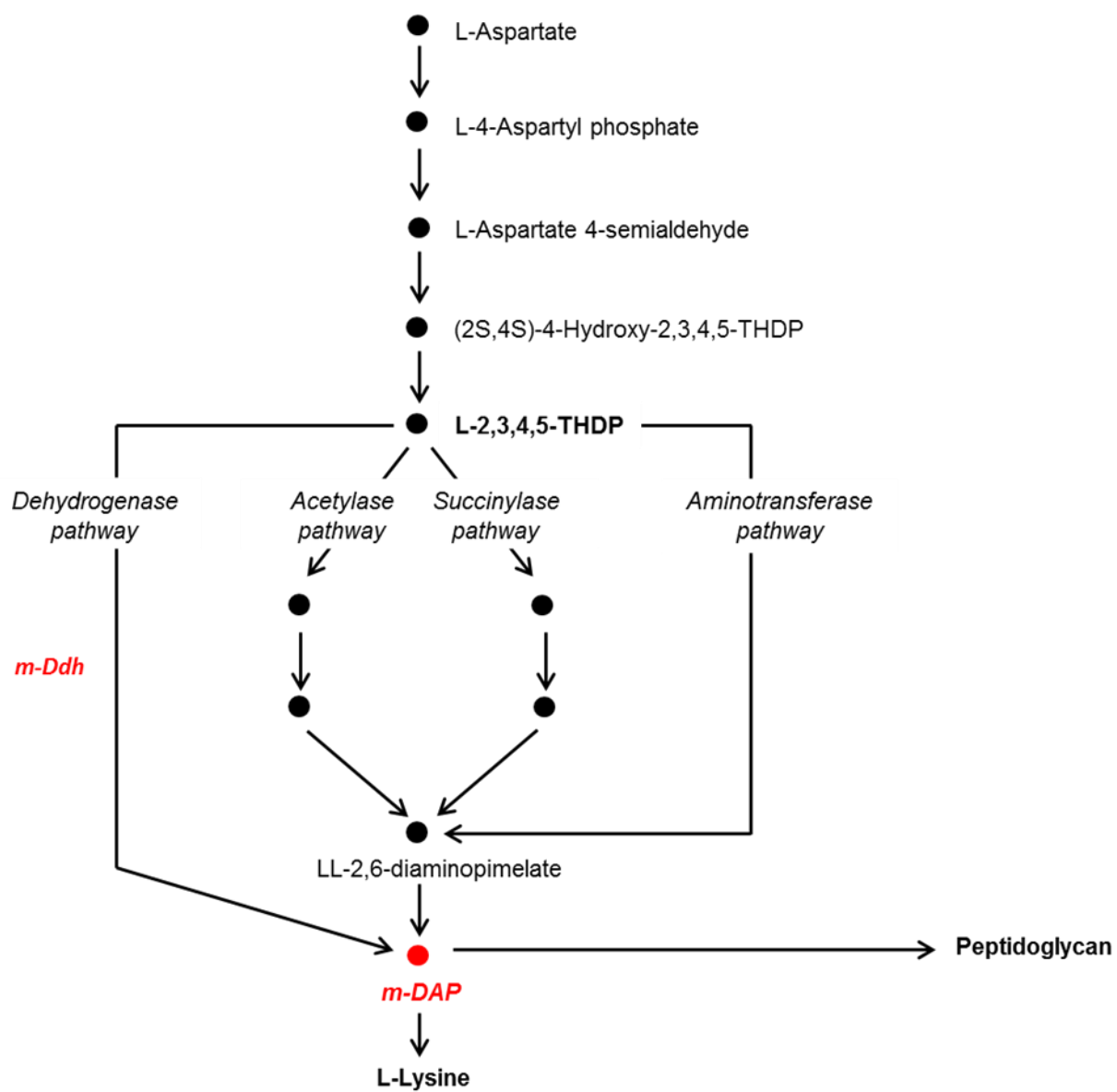


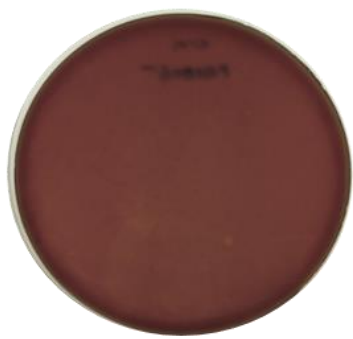
Figure 6. Variant pathways of lysine and meso-diaminopimelate biosynthesis.

Genes involved in the enzymatic reactions are omitted for general overview with the exception of the dehydrogenase pathway. Figure adapted from Born *et. al.*

experimentally true. To verify *m*-Ddh was essential in *P. gingivalis* strain W83, we knocked out the gene by transforming cells with a recombinant PCR product carrying an erythromycin resistance cassette (Erm^R) allowing for allelic replacement mutagenesis within the genome. Through this method, the antibiotic resistance cassette replaces the target gene and, if essential, results in non-viable cells following transformation. The disruption of the PG0806 gene with the Erm^R cassette resulted in no colony formation following electroporation and recovery in selective media (Fig. 7a). This result was repeated independently; suggesting that deletion of PG0806 is lethal to *P. gingivalis* and therefore essential. To show that the lack of colony formation was in response to the essentiality of the gene and not problems with the transformation, we simultaneously carried out allelic replacement mutagenesis for a non-essential hypothetical membrane protein (GenBank ID: AAQ65282.1). For this control we were able to obtain substantial colony formation (Fig. 7b).

Another component for a potential target is “druggability” or the chance a small-molecule will bind and have a significant effect on the protein’s activity [187]. Druggability can be assessed in several methods, e.g., proof-of-concept in similar proteins, conserved or targetable sequence motifs and structural analysis [188]. *m*-Ddh enzymes from several organisms referenced in UniprotKB/Swiss-Prot with known or binding sites predicted by similarity were aligned (Fig. 8). *P. gingivalis* has approximately 30% sequence identity to *C. glutamicum*, *L. sphaericus*, *C. thermocellum* and *U. thermosphaericus* and 70% sequence identity to *B. fragilis*. When analyzing the sequences of *m*-Ddh we found that the binding pocket for *m*-DAP was highly conserved. Druggable proteins have been shown to consist of a higher ratio of non-polar to polar amino acid residues (*m*-Ddh; 59.1% vs 40.9%) and a lower isoelectric point

a



b

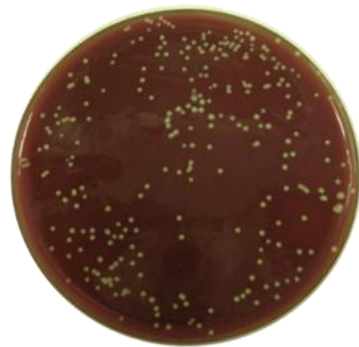


Figure 7. Allelic replacement mutagenesis for predicted essential gene target, *m*-Ddh.

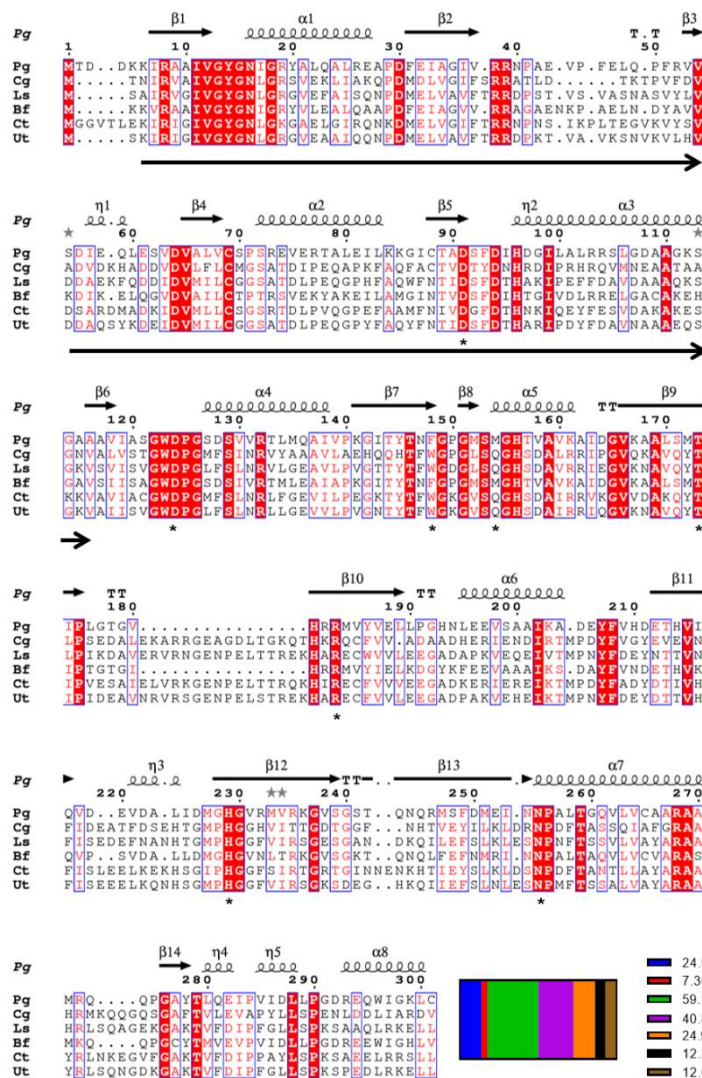
(**a**) Predicted essential gene PG0806 was transformed and plated on selective media resulting in no colony formation compared to (**b**) a non-essential gene control, validating the prediction that *m*-Ddh was essential.

(*m*-Ddh: 5.39 vs 7.44 in non-targets). Analysis showed that *m*-Ddh contains an oxidoreductase domain, a favored targeted enzyme class [189] and which has been previously determined with high confidence to be “druggable” by the Druggable Cavity Directory. Structural analysis by SYBYL revealed a solvent inaccessible binding cavity with residues corresponding to the *m*-DAP sequence alignment and the conserved motifs (Fig. 9). This binding cavity consists of a relatively deep and hydrophobic region which should consist of hydrophobic amino acid residues such as methionine, tryptophan and phenylalanine (Fig. 9). In addition, *m*-Ddh from *C. glutamicum* has been co-crystallized in a complex with the endogenous substrate [190]. Therefore, *m*-Ddh structure appears to be “druggable”, making it a suitable target for drug discovery.

Discussion

We aimed to present a quick and efficient manner to identify putative essential genes in bacterial species lacking genome-wide experimental data. This would be of great benefit, especially during drug discovery for emerging and re-emerging pathogens. Essential genes present novel targets for overcoming antibiotic resistance and developing new antimicrobial drug therapies. Genome-wide experimental efforts can be expensive and time-consuming, which has resulted in an increase in predictive methodologies. We established a framework for a novel approach derived from previous studies in a Gram-positive bacterium *S. sanguinis* [139]. We showed that essential genes can be linked to pathways related to three basic biological categories allowing for the prediction of essential genes based on gene function and annotation, subverting the need for orthologous genes comparison.

a



b

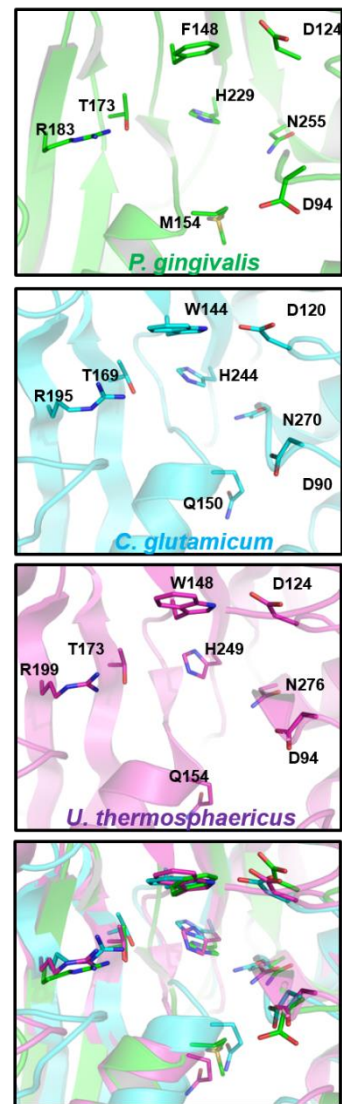


Figure 8. Sequence analysis of *m*-Ddh.

(a) Sequence alignment of *m*-Ddh from other bacterial organisms. The multiple sequence alignment analyses were performed using the T-Coffee multiple alignment. Alignment figure was generated in ESript 3.0. The putative binding sites of *Corynebacterium glutamicum* (Cg), *Lysinibacillus sphaericus* (Ls), *Bacteroides fragilis* (Bf), *Clostridium thermocellum* (Ct) and *Ureibacillus thermosphaericus* (Ut) cited in the sequence annotations in UniProtKB/Swiss-Prot and *P. gingivalis* (Pg) predicted based on homology are indicated by asterisks. The oxidoreductase domain for *P. gingivalis* is indicated by arrows below sequence. Secondary structure for *P. gingivalis* is annotated above the sequence. Relative percentage of characterized amino acid residues are shown below. (b) Secondary structure alignment of *m*-Ddh's putative binding site from *P. gingivalis* (green), *C. glutamicum* (cyan) and *U. thermosphaericus* (purple). Key residues are labeled, side chains are displayed as sticks and colored corresponding to atom type. Hydrogens were omitted for clarity.

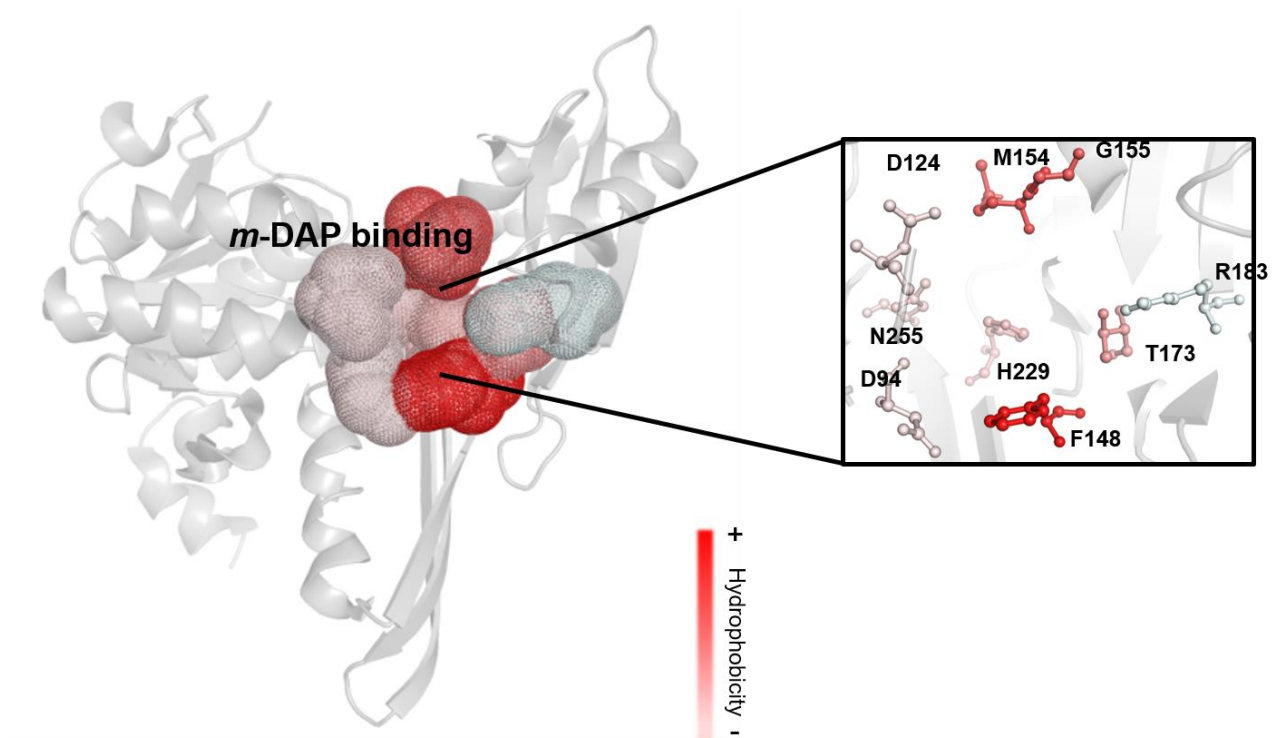


Figure 9. 3D structural analysis of *m*-Ddh from *P. gingivalis*.

Ribbon based structure of *m*-Ddh with the *m*-DAP binding domain. *m*-DAP binding cavity is displayed by wireframe surface and zoomed in view of cavity with corresponding a.a. residues labeled. Cavity colored by hydrophobicity (red = hydrophobic) is shown below.

In this study, we assessed the efficacy of our novel approach. By applying our gene annotation prediction model to the Gram-negative periodontal pathogen, *P. gingivalis*, and utilizing experimental data from a genome-wide transposon mutagenesis study performed in strain W83 of *P. gingivalis*, we showed that: 1) our assessment of essential pathways and knowledge of key genes within those pathways is accurate; 2) our model can easily be applied to a variety of bacterial species; 3) our prediction model may identify a core set of essential genes and provide important genomic data to refine our prediction model through the comparison of experimental data; 4) the benefit of having a bioinformatics/predictive approach together with an experimental approach; and 5) by applying our prediction model, we can accurately select a potential antimicrobial drug target.

From 1909 protein encoding genes annotated in KEGG, we predicted 212 essential genes. These genes fit within the three assigned biological categories (maintenance of the cell envelope, energy production and processing of genetic information) reaffirming the backbone of our model of essential genes. Having the advantage of recent experimental data by Klein *et al.*, we compared the two studies. The initial mutagenesis study was done in strain ATCC 33277 where 463 essential genes were identified. However, a new transposon mutagenesis study was carried out in strain W83 for the direct comparison of essential genes. In this experimental study, 759 putative essential genes were identified and a comparison resulted in roughly an 81% match to our model. It is important to note that the experimental data for strain W83 was not as comprehensive as the study from strain ATCC 33277, which is why not all the genes within the genome are accounted for. We also focused on what was determined to be putatively non-essential, to determine if we had false-positive predictions. From the study, 680 genes were determined to be non-essential. Our model was also accurate in not predicting non-essential

genes as essential. Only 7 genes were predicted from our model, but the majority (3 out of 5) had isozymes or paralogs that were determined to be essential.

Several limitations of our essential gene prediction model should be noted. First, our model does not identify hypothetical proteins as essential. Although hypothetical proteins may present a subset of undiscovered gene targets, it is unlikely that all of these genes would encode functions essential to the cell. During our experimental screen of *S. sanguinis*, out of 218 essential genes, only three were annotated as hypothetical. We also found that three of the hypothetical proteins shown to be essential during the transposon mutagenesis should not be expressed under the experimental conditions [191]. Genes not expressed may lead to false positives in transposon or allelic replacement mutagenesis studies. Second, genes whose functions are incorrectly annotated or incomplete will decrease the accuracy of our prediction. This would lead to genes falsely predicted to be essential or missed during our analysis completely. Organisms whose genome annotation is incomplete would also result in missed prediction during our analysis. For example, the KEGG genome annotation of *P. gingivalis* W83 is missing key enzymes in glycolysis for the conversion of PEP to pyruvate and pyruvate to acetyl-CoA. Another issue would be isozymes and/or paralogs. As our model is essentially predicting single gene knockouts, it would be difficult to determine whether one or both genes would be essential without experimental screening. Essential genes are predicted based on their biological function in regards to essential pathways. Ubiquitous genes or genes whose function is not linked to a specific pathway may be difficult to predict or may be missed. This is especially true for genes involved in genetic information processing. Many of these genes are involved in a variety of processes not defined to a specific biological functions. Lastly, genes can rarely be classified as absolutely essential or non-essential as the definition greatly depends on the

specified set of conditions [192]. Gene inactivation may result in a slow growth phenotype or a cell with decreased fitness that will die over time. For instance, deletion of the pyrophosphatase *ppa* gene in *E.coli* is viable for several hours until cellular pyrophosphates become too high, stopping cell growth [193]. Our model also assumed *P. gingivalis* cells were grown during *in vitro* conditions with rich media. Certain genes corresponding to the formation of vitamins and co-factors were noted as non-essential as they would be provided to the cell by the medium. Changes in these conditions, such as cells grown with minimal media, would therefore require additional pre-cursor essential genes.

A considerable benefit to this essential gene prediction model is the ability to quickly identify a set of potential drug targets. Due to the ease of our prediction, this method can be used to compare data across different genomes. By understanding crucial biochemical pathways and products, we could use model to identify species-selective essential genes as drug targets such as with our comparison between *S. sanguinis* and *P. gingivalis*. By comparing alternative pathways or differences in essential components (e.g., LTA versus LPS), we can identify specified essential genes. Information garnered from the comparison of experimental data can provide insight for our predictive model such as strain-specific essential genes. We aimed to support this hypothesis by selecting *m*-Ddh as a species-selective target for *P. gingivalis*. When selecting a drug target, it must not only be essential for the survival of the pathogen or disease virulence, it must possess certain sequence and structural features and three, have assayable activity. As a potential drug target, we were able to show through allelic replacement mutagenesis that *m*-Ddh is an essential enzyme. Previous studies comparing sequence and structural data between targets and non-targets showed druggable proteins are more likely to be of certain enzyme classes, contain more non-polar amino acids and have a lower isoelectric point, indicating molecules

more acidic in nature [189]. Based on these factors, *m*-Ddh is a “druggable” enzyme with a sequence and structural motif that has the ability to be targeted by small-molecules.

In conclusion, we have developed an accurate method to quickly predict a core set of essential genes through the analysis of gene function. While currently our data may not be comprehensive, our results demonstrate that this method can accurately be applied to organisms lacking experimental data. As genomic data and gene annotation becomes more complete our prediction model will also, and we anticipate more biological functions of hypothetical genes will be discovered with the advances of genome research. Further experimental data can also narrow the function of those experimental essential hypothetical genes down to the three basic biological categories that will increase discovery of gene functions. Additionally, more extensive studies can refine our prediction of key differences in essential genes between organisms. For example, many genes involved in fatty acid biosynthesis are essential for *Bacillus subtilis* and *Escherichia coli*, but the pathway is incomplete in *Mycobacterium genitalium*, indicating the product can be provided or is not essential [157]. This will ultimately facilitate our knowledge of essential genes. Overall, our study verifies the validity of our previous findings and by applying this to *P. gingivalis* we were able to identify *m*-Ddh as an essential and “druggable” target for the treatment of periodontal disease.

Chapter Two

Combinational Computer-based Drug Discovery to Identify Small-Molecule Inhibitors against *meso*-diaminopimelate dehydrogenase

Victoria N. Stone, Hardik I. Parikh, Glen E. Kellogg and Ping Xu

Victoria N. Stone performed the computational screening, the analysis of data and the preparation of the following manuscript. Dr. Hardik I. Parikh, Dr. Ping Xu and Dr. Glen E. Kellogg acted as advisors.

Background

Historically, antimicrobial drug discovery involved high-throughput screening of thousands of compounds for inhibitory activity, followed by secondary assays to identify the mechanism of action [97, 194]. Unfortunately, this was time consuming and costly, requiring on average 15 years and \$800 million [96, 195]. Coupled with a low success rate, this investment prompted a decline in antibiotic research by pharmaceutical companies, leading to a lack of new and effective drugs entering the market [126, 194]. However, with resistance and re-emerging pathogens on the rise, new approaches in drug development are critical for the future of antibiotics.

Recent advances in genomics, structural biology and computational chemistry have provided alternative strategies to traditional methods, giving rise to rational drug design [196]. The principle of rational drug discovery involves identifying or designing compounds based on the target's biological features. Consequently, the validation of a potential target is an important first step. We have previously demonstrated the understanding of essential gene functions can allow for the rapid prediction of essential genes as potential antimicrobial targets [139]. This understanding coupled with the knowledge of alternative pathways and differing metabolic requirements can be used to identify unique or species-limited gene targets. With a validated target, computer-based drug discovery (CBDD) can be a rapid, efficient and inexpensive way to identify and obtain a selection of potential antimicrobial compounds.

The use of a combinational CBDD approach (i.e., pharmacophore model, structure-based virtual screening (SBVS) and molecular docking/scoring) has been shown to be advantageous to the drug discovery process in many systemic disease studies, increasing their hit and success rate [197-199]. Due to the increasing access of essential genes (i.e. the DEG) [163] and protein

crystal structure data (i.e. PDB) [200] it has recently begun to be applied to antimicrobial research. A previous study in *M. tuberculosis* strain H37Rv successfully employed a multi-step screening strategy, combining modeling, SBVS and docking to identify a potent inhibitor against MtCM, a chorismate mutase [201]. Using a pharmacophore model we can generate a virtual description of molecular features essential to the desired protein-ligand interaction [202] based on observed experimental or *in silico* interactions [203]. This pharmacophore model can then be used for SBVS, which can replace costly and time-consuming experimental HTS assays. For SBVS a database with the structural data of millions of compounds is screened against the target protein to identify those that fit within the defined features and molecular docking is used to calculate and rank the binding mode for the select ligands within a defined region. Docking and scoring is a crucial step in CBDD. By defining an ideal pose, docking can aid in minimizing false positives prior to experimental studies and later can be integral in structure optimization [202]. However, evaluation of the optimal docking pose can vary based on the specific program and algorithm. For example, Goldscore uses a force-field method (summation of van der Waals and electrostatic interactions) [204] compared to CHEMPLP which utilizes empirical free energy scoring (enumeration of various types of interactions) [205]. To minimize the potential bias, studies suggest rescoring with different algorithms to determine the optimal binding confirmation [202]. As an additional means to evaluate the protein-ligand interaction, the free energy for optimal binding position can be predicted. HINT (Hydrophobic INTERactions), developed by Kellogg and Abraham [206], is an empirical force field based on experimental measurements of the small molecule partition coefficient, $\text{Log } P_{o/w}$. Since $\text{Log } P_{o/w}$ is thermodynamically related to free energy, the HINT score corresponds to the free energy of

intermolecular binding interactions. The combination of these approaches can provide a promising start for a novel antimicrobial drug discovery project.

We aimed to use periodontal disease and the keystone pathogen, *P. gingivalis* as an exploratory model for pathogen-specific drug targeting. We identified *m*-Ddh from *P. gingivalis* as a species-selective, essential and druggable target. The *m*-Ddh crystal structure from *P. gingivalis* (PDB ID: 3BIO) [207] was determined as part of the Protein Structure Initiative in 2007 [207, 208] allowing us to apply a combinational CBDD strategy to identify small-molecule inhibitors. In this study, we generated a pharmacophore model based on the natural substrate and previous *m*-Ddh inhibitor studies from *B. subtilis* and *C. glutamicum* [209]. We go on to adopt a HTS virtual screening method utilizing the ZINC drug-like database of commercially available chemicals to identify small-molecule inhibitors. Finally, we docked and scored the HTS results to identify a subset of high-ranking compounds as the initial steps in developing a novel strategy for antibiotic drug discovery.

Materials and Methods

Molecular modeling.

Protein structure. The structure of *meso*-diaminopimelate dehydrogenase (oxidoreductase, Gfo/Idh/MocA family member) from *P. gingivalis* strain W83, was crystallized as part of the National Institute of Health-National Institute of General Medical Sciences (NIH-NIGMS) sponsored Protein Structure Initiative (<http://www.nigms.nih.gov/Initiatives/PSI/>) [210, 211] and was solved at a resolution of 1.80 Å. The crystal structure data was downloaded from the Protein Data Bank (PDB ID: 3BIO) [207] and applied in our study. The binding site was identified by sequence homology to the ortholog in *C. glutamicum*, whose crystal structure was

previously determined in a complex with the substrate, *meso*-diaminopimelate (*m*-DAP), in the binding pocket (PDB ID: 2DAP) [190]. Using Sybyl X.1 (Tripos, St. Louis, MO), the protein was prepared for virtual screening and docking studies by extracting water molecules and co-crystallized ligands and deleting one of the two monomers. The pKa values of the amino acid residues within the binding pocket were predicted and the appropriate ionization states were assigned in SYBYL for a pH 10.5 based on the conditions of the *in vitro* enzymatic experimental assay. Appropriate atom types were assigned, hydrogens were added and the protein was minimized with Sybyl's Tripos force field.

Structure-based virtual screening. Virtual screening was performed with the UNITY module within the Sybyl-X molecular modeling program. Unity uses a directed tweak algorithm [212] to simulate molecular flexibility while screening small molecules. The binding pocket of *m*-Ddh was used as the target site, by constructing a variety of queries based on the pocket's properties. Over 9 million small molecules were screened *in silico* from ZINC [213] drug-like databases (<http://www.zinc.docking.org>).

Molecular docking. Docking and two-step scoring was used to evaluate the results of virtual screening. By visually inspecting and filtering the UNITY hits, we selected the top 132 small-molecule compounds for further computational analysis. We used GOLD (Genetic Optimization Ligand Docking) docking program v5.2) [204], targeting the binding site of *m*-Ddh. A sphere with radius of 12 Å from Arg183 was set as the docking region. This allowed for the inclusion of all residues expected to be within the binding site. The protein model was prepared for docking as described above. Conformational flexibility was allowed for the small molecules. We allowed for 50 GA (Genetic Algorithms) runs with a distance of 1 Å between clusters. The 132 compounds selected from our virtual screening hits were docked by GOLD, ranked by Goldscore

and then re-ranked by the CHEMPLP as implemented in GOLD. All docked compounds were then scored in a second pass by HINT (Hydropathic INTERactions) [214]. The binding mode corresponding to the highest-ranking HINT score for each compound was chosen as the best and most likely conformation for that compound. From these 132 compounds, the top 30% of the best-scored, structurally diverse compounds as ranked by HINT were re-docked and minimized with 10,000 iterations within the *m*-Ddh binding site. Finally, out of forty top scored small molecules, 11 compounds were commercially available and were purchased for assay. All images were generated in Pymol (<http://www.pymol.org>).

Cloning, expression and purification of *m*-Ddh. The amino acid sequence of *m*-Ddh from *P. gingivalis* was codon-optimized for expression in *E. coli* cells, synthesized and cloned into a pUC57 vector by GenScript (Piscataway, NJ). To introduce the 6×-HIS tag to the C-terminal end of the gene, the plasmid was digested at *NdeI* and *NotI* restriction sites. The digested fragments were loaded onto a 1% agarose gel and purified using MinElute® Gel Extraction kit (Qiagen, Valencia, CA). The purified DNA insert was ligated into a *NdeI*- and *NotI*-digested pET-21a (+) vector (Merck Millipore, Billerica, MA) by T4 DNA ligase (New England Biolabs, Ipswich, MA), yielding the expression plasmid pET-Ddh. The plasmids containing the DNA construct were isolated using QIAprep® Spin Miniprep plasmid (Qiagen, Valencia, CA) and sequenced at VCU Nucleic Acids Research Facilities (Richmond, VA).

The pET-Ddh plasmid was introduced into *E. coli* BL21 (DE3) pLysS (BioLine, Taunton, MA) and grown overnight in auto-inducing media ZYP5052 containing 100 µg/ml ampicillin at 37 °C. For purification, cells were disrupted by Emulsiflex C3 high pressure emulsifier (Avestin, Ottawa, Canada). Soluble protein was collected and separated from cell debris by centrifugation (20,000 × *g* for 20 mins at 4 °C). The resulting supernatant was loaded

onto a NTA-Ni²⁺ affinity column (Qiagen) pre-equilibrated with running buffer (25 mM Tris, 300 mM NaCl, 10 mM imidazole, pH 8.0). Unbound protein was washed off with wash buffer (25 mM Tris, 300 mM NaCl, 10 mM imidazole, pH 8.0) and chelated protein was eluted off with elution buffer (25 mM Tris, 300 mM NaCl, 100 mM imidazole, pH 8.0). Protein concentration was calculated based off of absorbance at 280 nm.

Molecular mass analysis of purified *m*-Ddh. The enzyme was examined by 12.5% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) stained with Coomassie Blue G-250 (Bio-Rad, Hercules, CA). Samples were boiled in 2X Laemmli buffer (4% SDS, 10% β -mercaptoethanol, 20% glycerol, 0.125M Tris-HCl, 0.004% bromophenol blue).

The molecular weight of the native enzyme was determined by gel filtration chromatography. Gel filtration was carried out using an Äkta Pure Protein Purification System (GE Healthcare, USA) with a 1 ml injection loop. The column was calibrated using Gel filtration protein standards of molecular weights ranging from 12,000 – 200,000 Da: standards 1, β -amylase (200 kDa) and cytochrome c (12.4 kDa); standards 2, alcohol dehydrogenase (150 kDa) and carbonic anhydrase (29 kDa); standards 3, albumin (66 kDa). 500 μ l of a 3.2 mg/ml purified protein was run through a Superdex 75 10/300 GL column (GE Healthcare, USA) at an elution rate of .5 ml/min with a 25 mM Tris-HCl, 300 mM NaCl, 100 mM imidazole buffer (pH 7.5) elution buffer. Fractions of 1ml volume were collected. Fractions were then analysed by spectrophotometry. Protein size was calculated by standard curve of molecular mass vs. V_e (elution volume)/ V_o (void volume) for each protein standard.

***m*-Ddh kinetic assay.** The enzymatic activity for *m*-Ddh was determined by observing the standard oxidative deamination reaction of the substrate *meso*-diaminopimelate [215]. The reaction contained 400 μ M of *meso*-diaminopimelate (Sigma-Aldrich, St. Louis, MO), 180 μ M

NADP⁺ (Sigma-Aldrich, St. Louis, MO), 200 mM glycine-KCl-KOH buffer (pH 10.5), and the enzyme in a final volume of 1 ml. The reaction was initiated with the addition of NADP⁺. The reaction velocity was calculated from the increase in absorbance at 340 nm, spectrophotometrically monitored at 25 °C, where one unit of enzyme was defined as the amount of enzyme catalyzing the formation of 1 mmol of NADPH per min.

Determination of kinetic parameters. Initial velocity measurements for *m*-DAP and NADP⁺ were determined at 25 °C in a similar reaction for the standard oxidative deamination reaction assay. The reaction contained 200 mM glycine-KCl-KOH buffer (pH 10.5) with *m*-DAP as the variable substrate with concentrations between 0.001 mM and 1 mM and NADP⁺ held constant at a saturating concentration of 0.5 mM or NADP⁺ as the variable substrate with concentrations between 0.01 mM to 1 mM and *m*-DAP held constant at 0.5 mM. K_m and V_{max} values were determined through non-linear fitting. All assays were performed in triplicates and non-linear fitting Michaelis-Menten data were calculated from Graphpad Prism v5.04 (Graphpad, San Diego, CA).

Mutagenesis of *m*-Ddh. For site directed mutagenesis, pET-Ddh plasmid was used as a template for single residue site-directed mutagenesis of *m*-Ddh (R183Y, R183K, D124E and D124A). The phosphorylated primer set for each mutant was as follows. Forward R183Y: 5'- ACG GGT GTG CAT CGT TAT ATG GTC TAT GTG GAA-3' and Reverse R183Y: 5'- TTC CAC ATA GAC CAT ATA ACG ATG CAC ACC CGT-3'; Forward R183K: 5'- ACG GGT GTG CAT CGT AAA ATG GTC TAT GTG GAA -3' and Reverse R183K 5'- TTC CAC ATA GAC CAT TTT ACG ATG CAC ACC CGT-3'; Forward D124E: 5'- C GCA TCA GGC TGG GAA CCG GGT AGT GAT TCC-3' and Reverse D124E: 5'- GGA ATC ACT ACC CGG TTC CCA GCC TGA TGC G -3'; Forward D124A: 5'- ATC AGG CTG GGC GCC GGG TAG TG -3' and Reverse

D124A 5'- GCG ATA ACC GCA GCT GCA -3'. All PCR reactions were performed with an initial denature of 98 °C for 30 sec, 25 cycles of 98 °C for 10 sec, 56 °C for 30 sec, 72 °C for 2 min 45 sec and a final extension of 72 °C for 7 min. PCR products were digested with 1 U of *DpnI* (New England Biolabs, Ipswich, MA) at 37 °C for 1 h and inactivated for 20 min at 80 °C.

The plasmids containing the mutation were sequenced at VCU Nucleic Acids Research Facilities (Richmond, VA). The final product was used to transform into *E. coli* DH5 α T1 competent cells. Phusion High-Fidelity Taq DNA polymerase (New England Biolabs, Ipswich, MA) was used in all reactions. Mutants were transformed, expressed and purified as previously described for WT.

Results

***In silico* analysis of *m*-Ddh structure and function.** The enzyme catalyzes the NADP⁺ dependent oxidative deamination of *m*-DAP (Fig. 10). The reaction is reversible and the biologically favored reaction involves the NADPH dependent conversion of L-2,3,4,5-THDP to *m*-DAP by reductive animation. During this reaction L-2,3,4,5-THDP first has a spontaneous ring opening to produce L- α -amino- ϵ -ketopimelate. L- α -amino- ϵ -ketopimelate catalyzed by *m*-Ddh generates an imine intermediate before yielding *m*-DAP. Crystal structure data suggests the protein is similar to previously studied *m*-Ddhs [182, 216-218]. It is a homodimer and sequence analysis indicates it consists of three main domains: a NADP⁺-binding domain which corresponds to the conserved GXGXXG sequence found within the N-terminal $\beta\alpha\beta$ -protein fold [219], a dimerization domain and an oxidoreductase domain (Fig. 11). Studies in *C. glutamicum* and *B. sphearicus* show that *m*-Ddh has a strict specificity for the D-amino acid center of *m*-DAP being placed near the NADP(H) domain [220-222].

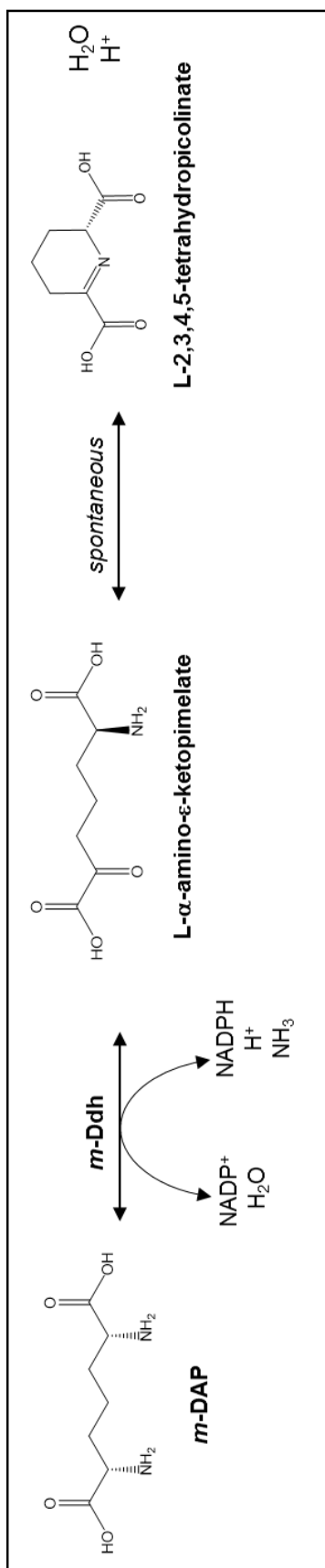


Figure 10. Schematic for the *m*-Ddh catalyzed biochemical reaction.

To assess the binding model of the natural substrate within the protein pocket we docked *m*-DAP (Fig. 12). The protein was prepared as described in “Materials and Methods” and we used GOLD docking program v5.2, selecting the 9 a.a. residues previously predicted to be the binding site as the area of the docking region. *m*-DAP binds within a deep groove situated so that the two carboxylate groups face the protein side allowing for hydrogen bond interactions (Arg183, Thr173, Met154 and Gly155). The two amine groups face away from the protein allowing for potential solvent interactions. The positioning places the D-amino acid center near the NADP⁺ binding site to allow for the hydrogen exchange and the enzymatic reaction to occur. This binding mode corresponded to what has been observed in crystal data studies [190, 220].

***In vitro* analysis of *m*-Ddh structure and function.** To first confirm the structural and functional *in silico* analysis of *m*-Ddh, we synthesized the gene and expressed it in *E. coli*. The gene sequence encoding for *P. gingivalis m*-Ddh was codon-optimized and cloned into a T7 pET-21a (+) expression vector carrying a C-terminal 6x HIS tag. The protein was isolated following the expression and purification described in “Materials and Methods”. The protein is 301 a.a. residues with a calculated molecular weight of 32 kDa, corresponding to the migration of the monomeric structure on SDS-PAGE (Appendix, Supp. Fig. 2a). The crystallized structure of *m*-Ddh from *P. gingivalis* indicated that the enzyme exists as a homodimer, the native molecular weight was determined by gel filtration chromatography where two prominent peaks eluted similarly to standard 3 (Appendix, Supp. Fig. 2b) at a calculated size of 66 kDa, demonstrating that the active enzyme exists as a dimer.

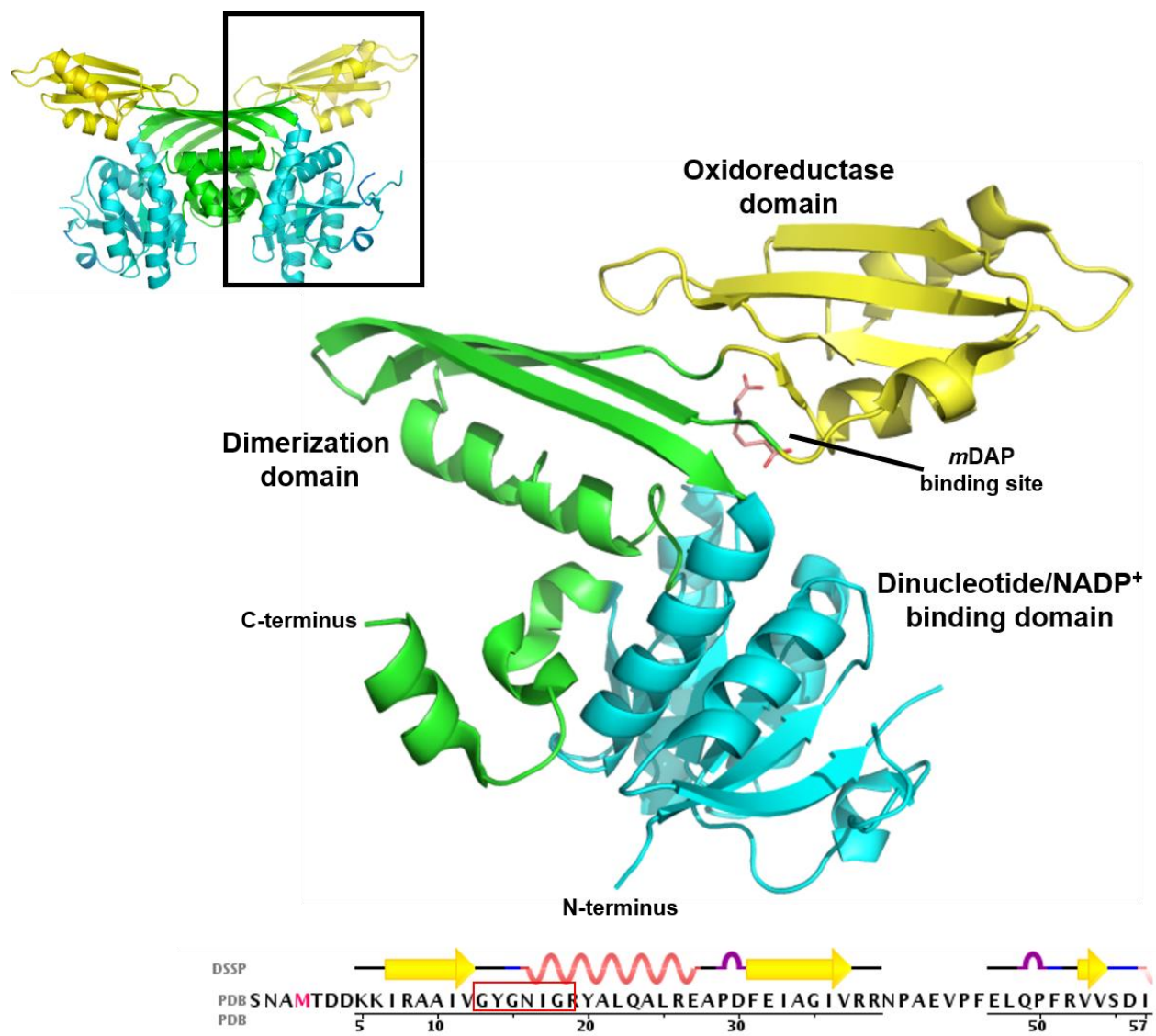


Figure 11. Structural domains of *m*-Ddh from *P. gingivalis*.

Dimer (top) and monomer (bottom) structure of *m*-Ddh from *P. gingivalis*. Structure consists of three main domains which are highlighted by different colors. The substrate or oxidoreductase domain (yellow), the dimerization domain (green) and the NAD(P) domain (blue). The NADP domain found in the N-terminal sequence of *m*-Ddh corresponds to the conserved GXGXXG sequence within a $\beta\alpha\beta$ structural motif. 3D structure of *m*-Ddh was generated in Pymol (<http://www.pymol.org>). *m*-DAP docked within the binding pocket is shown in pink.

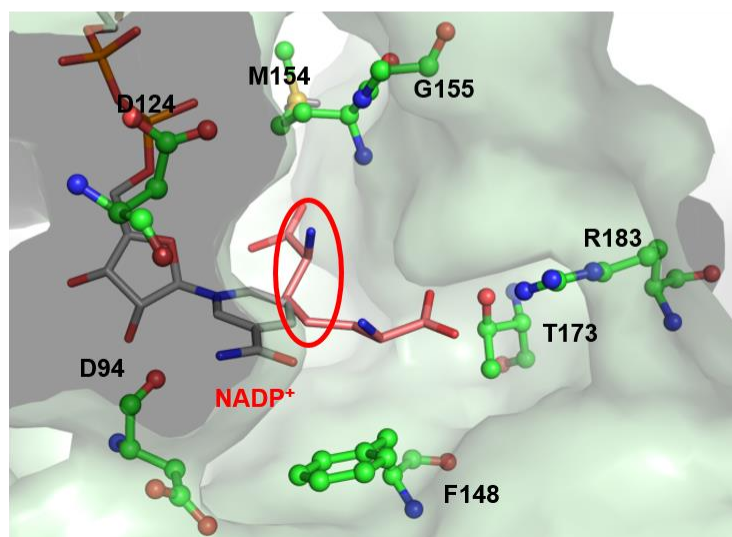
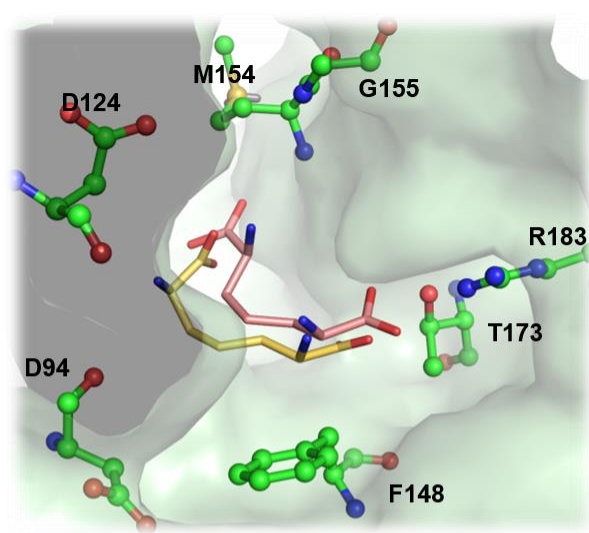
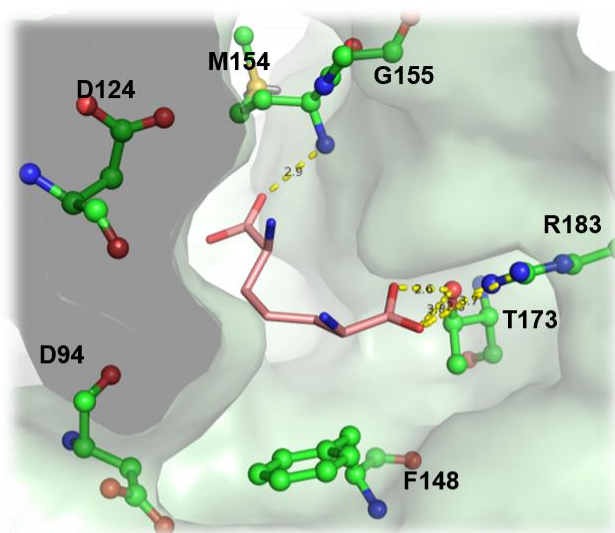
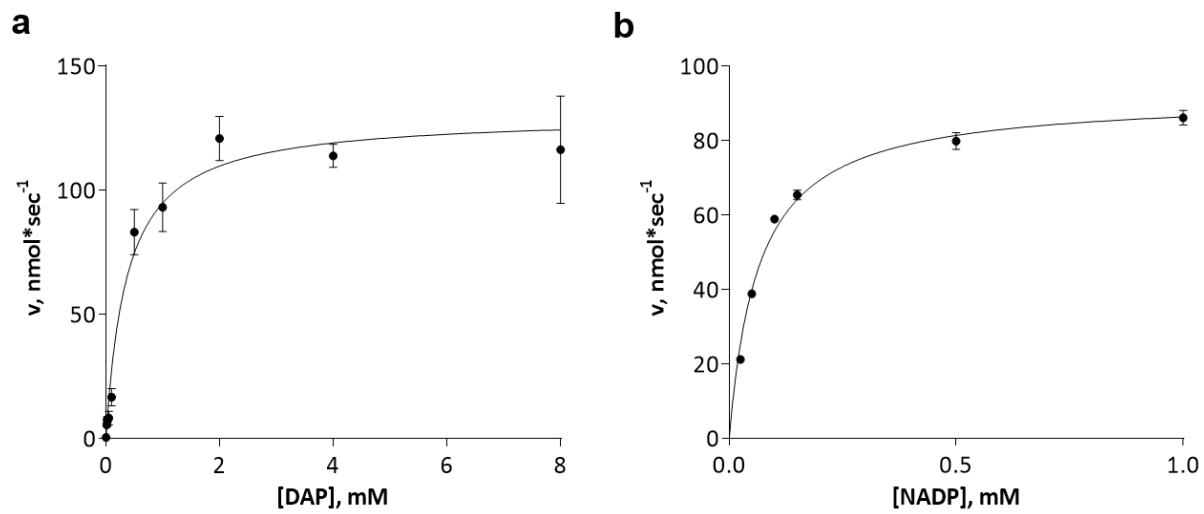


Figure 12. Molecular docking analysis of *m*-DAP binding model.

GOLD was used to dock the natural substrate, *m*-DAP into the predicted binding pocket of *m*-Ddh. The optimal binding model determined by HINT score is displayed above (pink). Binding position of *m*-DAP from previous crystal data studies (PDB ID: 2DAP) is overlaid with predicted docking model (yellow) to show similar binding. NADP⁺ docked in its binding pocket is shown below to display the positioning for the hydrogen exchange and the enzymatic reaction. Key residues for intermolecular interactions are labeled, displayed as ball and sticks and colored corresponding to atom type. Hydrogens were omitted for clarity. Potential hydrogen bonding interactions between *m*-Ddh and a.a. residues are shown by yellow dashed lines.

We next examined the Michaelis kinetics to evaluate the activity of the enzyme and determine the kinetic parameters for the oxidative deamination of *m*-DAP. This reaction is spectrophotometrically observable at 340 nm. This allowed for the development of a standard screening assay, monitoring the reduction of NADP⁺ to NADPH. Analysis of the initial velocity showed typical Michaelis-Menten kinetics (Fig. 13). The apparent K_m and V_{max} for *m*-DAP was determined to be 370 μ M and 130 nmol sec⁻¹ and 60 μ M and 92 nmol sec⁻¹ for NADP⁺ (Fig. 13) respectively.

Docking studies indicated that the carboxylate groups were essential for the binding through hydrogen bonding interaction. To verify this we selected Arg183 to mutate with a favored residue substitution or an un-favored substitution (Table 3; Fig. 14a). An Arg mutation to an un-favored substitution (R183Y) resulted in a significant decrease in substrate affinity (35-fold increase in the apparent K_m) compared to WT but retained similar catalytic efficiency. This was in contrast to the favored Arg substitution (R183K) which maintained similar K_m and K_{cat}/K_m values as the WT. Comparison of these two mutants with the WT docking showed a loss in a favorable hydrogen bonding between R183 and *m*-DAP during the Y183 mutation which led to a decrease in the calculated HINT interaction. Mutant R183K had a similar pose as WT, reflected by a strong HINT interaction between the carboxylate group of *m*-DAP and the amine group from K183. Asp124 is positioned within the protein for both NADP⁺ and *m*-DAP interaction. A favored mutation (D124E) resulted in a slight increase in substrate affinity (2.7-fold decrease in the apparent K_m) and an increase in catalytic efficiency. In contrast, the D124A mutation had a dramatic change in both the apparent K_m and the catalytic efficiency resulting in almost no measurable activity (Table 3; Fig. 14b). Docking showed similar potential hydrogen



Varied Substrate	K_m (μM) ^a	V_{\max} ($\text{nmol} \cdot \text{sec}^{-1}$) ^a
<i>m</i> -DAP	372 ± 0.05	130 ± 4
NADP ⁺	65 ± 0.004	92 ± 2

^a Mean \pm s.d.
n=3

Figure 13. Kinetic analysis of purified *m*-Ddh from *P. gingivalis* strain W83.

Characterization of kinetic properties of *m*-Ddh. Oxidative deamination reactions were performed in the presence of increasing concentrations of (a) *m*-DAP as the variable substrate with NADP⁺ fixed at a saturating concentration or (b) NADP⁺ as the variable substrate and *m*-DAP held constant. K_m and V_{max} values were determined through non-linear fitting.

Table 3. Kinetic analysis of *m*-Ddh mutants.

<i>m</i> -Ddh	K_m (μM)	K_{cat} (s^{-1})	K_{cat}/K_m ($\mu\text{M}^{-1} \text{s}^{-1}$)
WT	374.2	1.30×10^9	3.48×10^6
R183Y	12990	1.50×10^9	1.16×10^5
R183K	1109	1.92×10^9	1.73×10^6
D124E	139.9	8.97×10^9	6.41×10^7
D124A	ND	ND	ND

ND, not able to kinetically determine

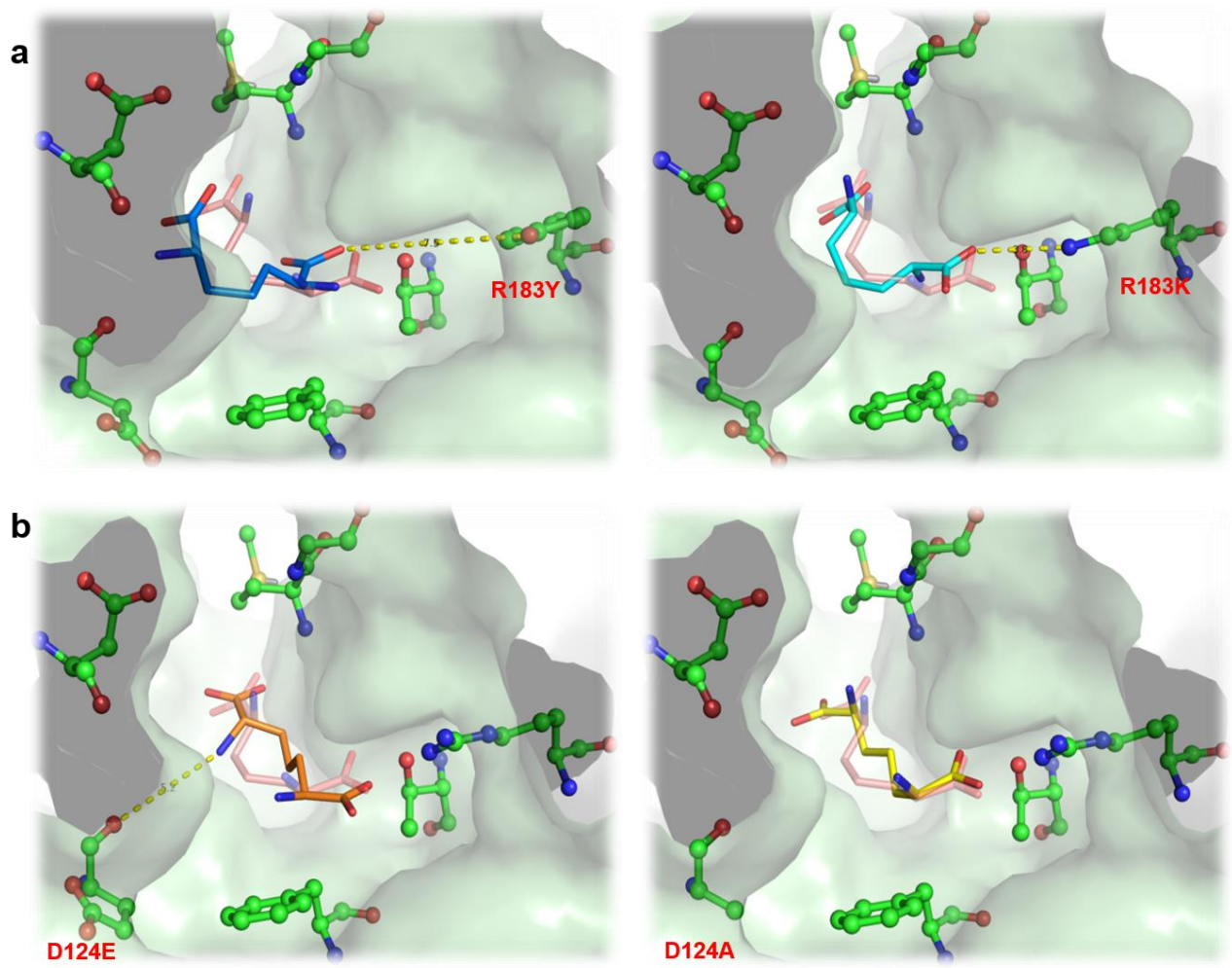


Figure 14. Site-directed mutagenesis of binding pocket.

GOLD was used to dock the natural substrate, *m*-DAP into *m*-Ddh with single a.a. substitutions of the predicted binding pocket. The optimal binding model is displayed above. **(a)** Arg183 mutated to R183Y left and R183K right with *m*-DAP docked. **(b)** Asp124 mutated to D124E left and D124A right with *m*-DAP docked. Key residues for intermolecular interactions are displayed as ball and sticks and colored corresponding to atom type. Mutated residues are labeled in red. The optimal *m*-DAP docked for WT *m*-Ddh is shown in pink to highlight changes in binding confirmation. Changes in potential hydrogen bonding interactions that would affect activity is displayed as a dashed yellow lines. Hydrogens were omitted for clarity.

bonding for the D124E mutant, however there was a more favorable acid-base interaction compared to WT. Conversely, D124A showed a slightly unfavorable base-base interaction between the carboxylate group of *m*-DAP and the backbone of Asp.

Generation of pharmacophore model. To identify a binding model for inhibition, we searched the literature for known inhibitors. Unsaturated analogues of *m*-DAP, containing a planar α -carbon and lacking the active D-amino acid amine center of *m*-DAP have been shown to be strong inhibitors of *m*-Ddh isolated from *Bacillus sphaericus* and *C. glutamicum* [209]. It was assumed the analogue inhibitors bind in manner opposite to that of the substrate; thus, the non-reactive L-amino acid center is positioned near the C-4 position of the co-substrate NADP⁺. This would prevent the oxidation reaction and hydride exchange that normally would occur between the substrate and co-substrate. We obtained two of these previously reported compounds; testing *in vitro* showed dose-dependent inhibition against *m*-DAP from *P. gingivalis*.

The X-ray crystal structure was modeled and the unsaturated analogue inhibitors (Compounds **1** – **3**) as well as the *m*-DAP substrate were docked into the *m*-Ddh binding pocket to identify the features that should be important for inhibitor interactions (Fig. 15a). The docking model which best fit the expected *in vitro* interaction and displayed high docking scores was used to generate a pharmacophore model. Based on the best ranking interaction, the pharmacophore model focused on four features that were shared between the inhibitors and substrate: 1) a hydrophobic region complementary to amino acid residues Trp123 and Phe148; 2) a ligand donor atom complementary to residues Asp94 and Asp124; 3) a negative (acceptor) center complementary to the side chain of residue Ser153 and the backbone of residues Met154 and Gly155; and 4) a negative (acceptor) center complementary to the side chains of residues

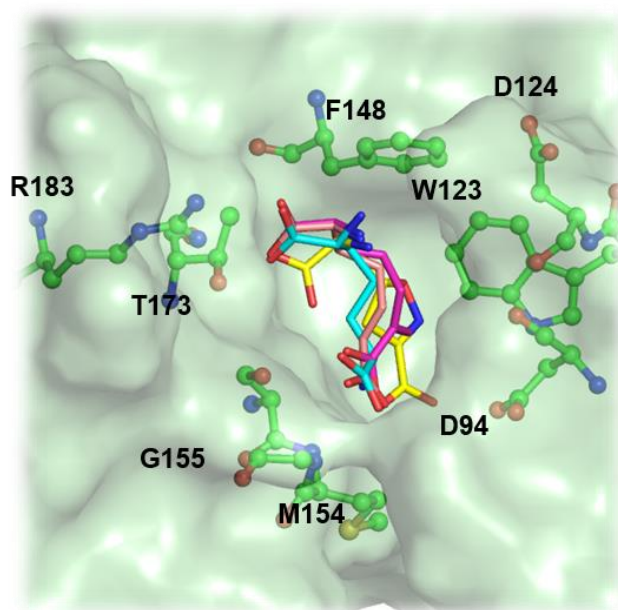
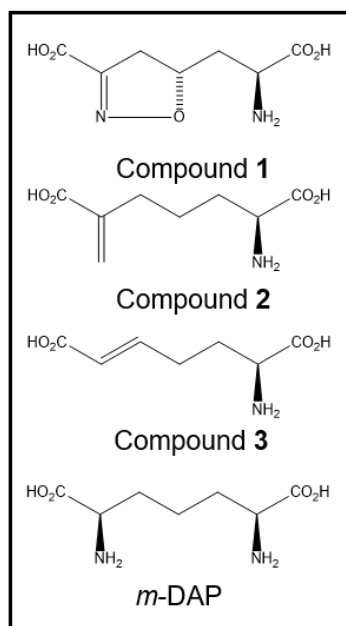
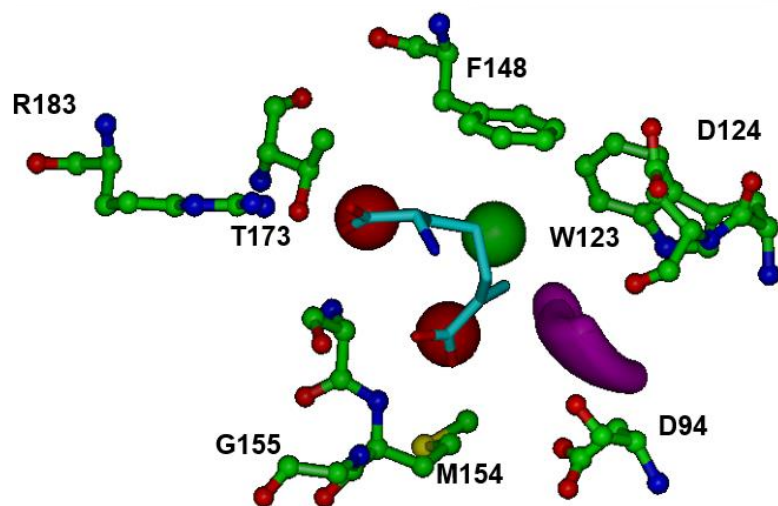
a**b**

Figure 15. Generation of pharmacophore model.

(a) Structure of *m*-DAP and inhibitor analogs that were previously shown to be active against *m*-Ddh in *C. glutamicum* and *B. sphaericus*. (b) Compounds docked into *m*-Ddh binding site and conserved interactions were identified. (c) Pharmacophore model with selected core features for inhibitor identification during virtual screen. The model focused on four features: first, a hydrophobic region complementary to amino acid residues Trp123 and Phe148 (green), second, a ligand donor atom complementary to residues Asp94 and Asp124 (purple), third, a negative center complementary to the side chain of residue Ser153 and the backbone of residues Met154 and Gly155 (red) and fourth, a negative center complementary to the side chain of residues Arg183 and Thr173 (red). The interaction was also restricted for an area 12Å in distance for Arg183. Key residues are labeled, displayed as ball and sticks and colored corresponding to atom type. Hydrogens were omitted for clarity.

Arg183 and Thr173. The interaction was also restricted to a sphere of radius 12 Å centered around Arg183. This was because from our docking model Arg183 was seen to form hydrogen bonds with the carboxylate groups of the substrate analogues and site-directed mutagenesis decreased the substrate-protein binding affinity by 32-fold.

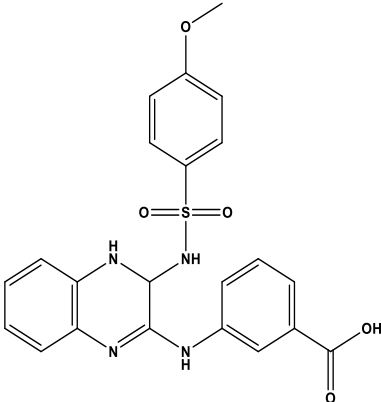
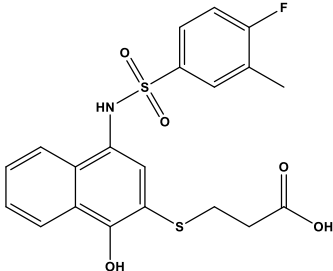
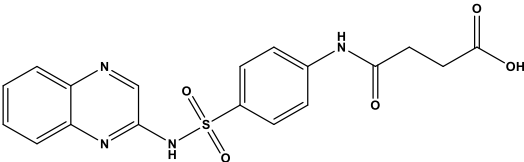
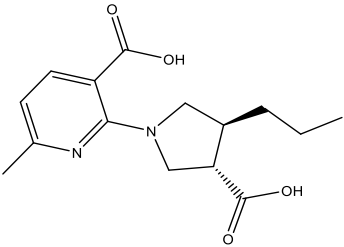
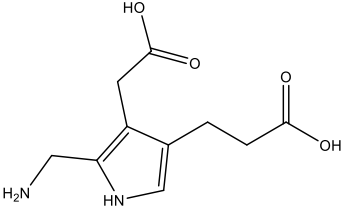
High-throughput virtual screening for identification of small-molecule inhibitors. The pharmacophore model shown in Figure 15b was used in a high-throughput virtual screen of the ZINC 3D database to identify small-molecule inhibitors that would fit the query constructed from the pharmacophore. ZINC (Zinc Is Not Commercial) is a publicly available listing of molecules that are reportedly available for purchase, organized in a manner appropriate for virtual screening studies. In simple terms, a compound was classified as a hit if it fit all of the features defined as mandatory in the model. The screening of more than 9 million compounds within the ZINC database resulted in more than several hundred hits. Since the goal of virtual screening is to identify unique compounds and scaffolds that have the potential to be developed into active inhibitors, a filter was applied to remove compounds within the hit list too structurally similar to one another. The resulting list was then filtered for drug likeness (i.e., with algorithms based on Lipinski's Rule of Five [188]) to remove compounds and scaffolds that were unlikely to have reasonable physiochemical properties.

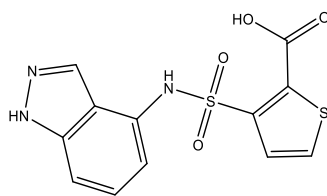
Molecular docking and scoring of small-molecule inhibitors. Compounds passing through these filters were then docked to the binding site of *m*-Ddh with GOLD (Genetic Optimization of Ligand Docking) program to predict their binding affinities and to assess the modeled compound-protein interactions. Prediction of the best fit binding model of each compound, again within 12 Å of Arg183, was determined and scored by Goldscore. The models with top docking scores were re-docked to the binding site with the same docking parameters and rescored by

CHEMPLP. A filter based on binding pose was applied and compounds that interacted favorably, mostly via hydrogen bond, with the key residue Arg183 were identified. This filter yielded 132 hits, which were then scored by the HINT force field. The binding mode corresponding to the highest HINT score for each compound was then re-docked and minimized within the *m*-Ddh binding site. From these 132 compounds, the top 30% of the best HINT-scored, structurally diverse compounds were set-aside as the 48 final hits. Finally, samples of the commercially available compounds in this group were purchased for future screening assays. The HINT scores and compound structures of each of these 11 compounds (**4** – **14**) are shown in Table 4.

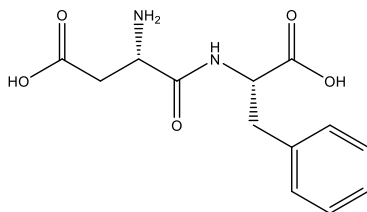
***In silico* analysis of binding confirmations.** The optimal binding pose for the six top ranking compounds determined by HINT score are shown in Figure 16. The protein-inhibitor interaction involved potential hydrogen bonding interactions with the backbone for Met154, Ser153 and Gly155. The majority of the compounds possessed aromatic structures forming favorable hydrophobic interactions with Phe148 or Trp123. The sulfonamide functional group on compound **4**, **5** and **6** show potential hydrogen bonds with Arg183 and Thr173. Due to distance few compounds show potential hydrogen bonding with the Asp functional group. The exception was compound **10** which lost hydrogen bonding between Arg183 and Thr173 to gain favorable hydrogen bonding interactions with Asp93. Overall, these compounds maintained the desired interaction determined through the pharmacophore model.

Table 4. 2D structures and scoring of top-ranking compounds

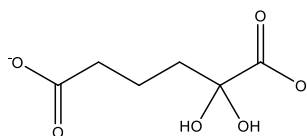
	Structure	HINT
Compound 4		3112
Compound 5		3014
Compound 6		2876
Compound 7		2095
Compound 8		4379

Compound **9**

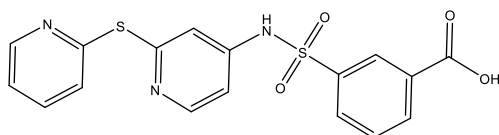
2569

Compound **10**

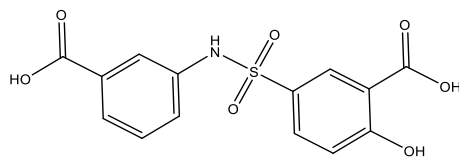
4274

Compound **11**

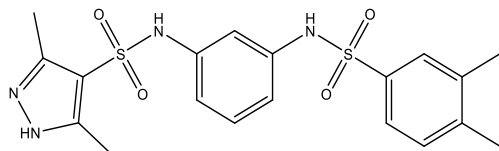
2344

Compound **12**

1989

Compound **13**

4190

Compound **14**

1692

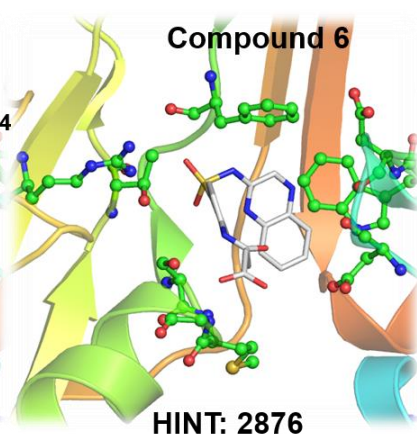
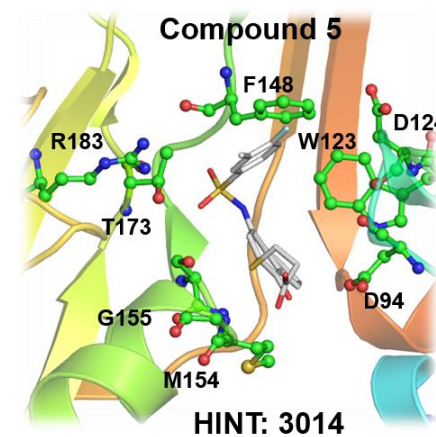
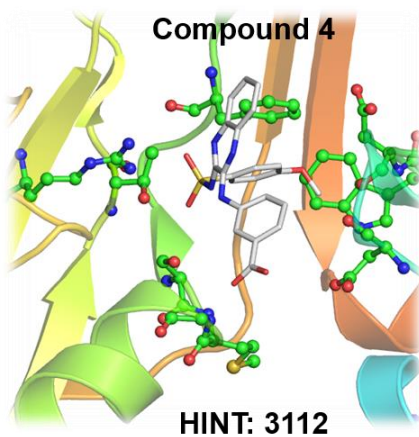
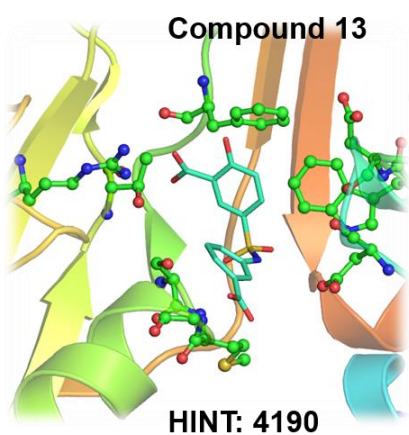
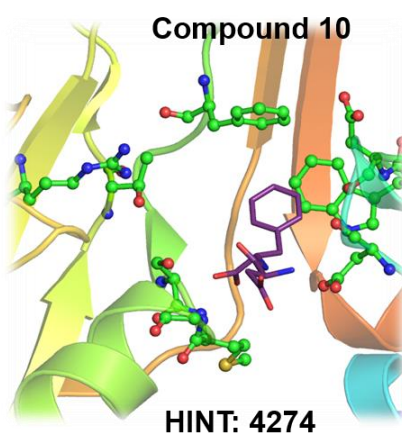
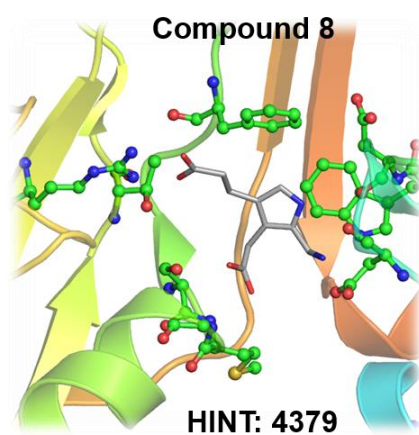


Figure 16. Optimal binding mode for top-ranking compounds.

Optimal binding model for the top-ranking compounds determined through HINT are shown above (see Table 4). Key residues are displayed as ball and sticks and colored corresponding to atom type. Residues labeled in the bottom middle figure correspond to all figures. Hydrogens were omitted for clarity.

Discussion

Novel therapeutics have not kept paced with the need to overcome antibiotic resistance and fight new and re-emerging pathogens. Traditional methods principally relied on empirical screening, screening large numbers of compounds for whole-cell activity, flushing out the mechanism of action and verifying the feasibility of the target later. This resulted in a costly and time-consuming endeavor that was especially discouraging for small startups and academia which lack the resources of larger pharmaceutical companies. In addition, as the ‘golden era’ of antibiotic discovery waned, this approach proved ineffective. However, steady advances in computational chemistry, protein chemistry and genomics have resulted in successful drug discovery through CBDD for systemic diseases such as cancer [202]. Utilizing a combination of CBDD techniques, a better understanding of a target can be gained, the properties essential for the activity or inhibition assessed and a diverse set of inhibitors identified.

In this study, we aimed to present a rationalized approach to antibiotic drug discovery. We employed a combinational CBDD approach which incorporated pharmacophore models, SBVS and molecular docking. Ultimately screening more than 9 million small-molecule compounds to identify potential inhibitors against *P. gingivalis* and utilize a novel approach in antimicrobial drug discovery. By applying this method to the Gram-negative periodontal pathogen, *P. gingivalis*, we showed that: 1) we can accurately assess the structure/function of our target through computer-based molecular modeling; 2) we can identify key molecular features for inhibitor screening; and 3) utilizing several CBDD techniques in parallel can lead to a rapid selection of inhibitors.

The crystal structure data of *m*-Ddh was used to first examine the structure and related function of the enzyme. The enzyme is an active homodimer that binds *m*-DAP and NADP⁺

within close proximity at the N-terminus of the protein. It has been shown to have strict specificity for the substrate which must bind in a certain orientation. This would allow for specific hydrogen bonding interactions with the carboxylate groups of *m*-DAP. It also positions the catalytically active amine group near the NADP⁺ moiety. As *m*-Ddh from *P. gingivalis* has not been characterized *in vitro*, we expressed and purified the protein from *E. coli*. Gel filtration analysis confirmed the assumed tertiary confirmation consistent with other characterized *m*-Ddhs. Previous studies of *m*-Ddh and its role in lysine biosynthesis have focused on the enzyme from *Corynebacterium* [190], *Bacillus* [184], and *Ureibacillus* [217]. However, prior to our study there was no enzymatic data on *m*-Ddh in *P. gingivalis* available. Therefore, we determined the kinetic properties through the assessment of the enzyme assay. Analysis of the kinetic data showed a K_m value of 370 μM with a V_{\max} of 130.1 nmol sec^{-1} for the substrate *m*-DAP and a K_m of 60 μM with a V_{\max} of 91.95 nmol sec^{-1} for NADP⁺. We next used site-directed mutagenesis to further evaluate the *in silico* interaction. The side chain of Arg183 was shown to form hydrogen bonds within the active site. Mutation of Arg183 to Tyr was still catalytically active but showed a large decrease in substrate binding compared to the WT. This may indicate the replacement of Arg to the aromatic structure of Tyr may prohibit *m*-DAP access into the full length of the active site but still allows for the enzymatic reaction near the D-amino center. D124 is found within the overlap between the substrate and co-substrate binding sites. Asp residues within active sites are known to form stabilizing hydrogen bonds through salt bridges [223]. The Ala substitution left the enzyme almost inactive. However a Glu substitution, which is similar to Asp in structure and function, slightly increased the binding and efficiency. This could indicate D124 may play an essential role in *m*-Ddh protein stabilization.

By analyzing the binding model of substrate analogue inhibitors we were able to define a pharmacophore model with consistent features between three inhibitors and the substrate. Our four key features mainly focused on the potential hydrogen bonding interactions. First identified was a hydrophobic region from the aromatic residues Trp123 and Phe148. This interaction would allow for another ring within that region and lead to potential hydrophobic stacking interactions. The second feature, a ligand donor atom complementary to residues Asp94 and Asp124 would allow for hydrogen bonding between the carbonyl oxygen groups. The third feature is a negative (acceptor) center complementary to the side chain of residue Ser153 and the backbone of residues Met154 and Gly155. The backbone of these residues would allow for potential hydrogen bonding with the ligand acting as a hydrogen acceptor. The last feature was a negative (acceptor) center complementary to the side chains of residues Arg183 and Thr173. This feature is similar to feature three but involves the side chains for potential bonding. This characterization allowed for the application of a rational HTS strategy utilizing the ZINC compound database. Compounds identified were then subjected to docking analysis through GOLD and two distinct scoring approaches. This resulted in more than 100 compounds. Finally a subset of the commercially available compounds which were computationally determined to be the best fit were selected for future *in vitro* analysis.

In conclusion, we have successfully applied a CBDD method to identify a select group of structurally diverse small-molecule compounds against *m*-Ddh. While currently this data is preliminary, our results demonstrate that this multifactorial method can accurately be applied to antimicrobial drug discovery and due to the feasibility, can be used across a variety of infectious pathogens. In addition, the rational selection and preliminary *in silico* screening should lead to a higher success rate than traditional trial and error experimental screening. This study may also

present a better understanding of our validated target. Through computational modeling we can further study the structure of the protein. This is important, as molecular features correspond to the activity of the protein and by understanding these features a compound can then be selected or designed to interact against these features. Ultimately, this may result in a more in-depth pharmacophore model, improved HTS drug screening or structural optimization of active inhibitors. Overall, our study presents the rationale for this approach and by applying this to *P. gingivalis* we were able to identify potential inhibitors for the treatment of periodontal disease. Coupled with further experimental data we can verify our computational analysis and this should aid in the progression of an antimicrobial drug discovery model.

Chapter Three

***In vitro* Characterization of Small-Molecule Inhibitors against *meso*-diaminopimelate dehydrogenase**

Victoria N. Stone and Ping Xu

Victoria N. Stone performed the experiments, the analysis of data and the preparation of the following manuscript. Dr. Ping Xu acted as an advisor.

Background

Undoubtedly, CBDD is an exceptional way to identify a set of inhibitors against a specific target. The rational assessment of potential interactions allows for the selection of small molecules with a higher probability of binding, thereby increasing the efficiency of HTS [202]. However, their activity is based on computational algorithms and is therefore theoretical. It does not provide information on the potency, pharmacokinetics or cytotoxicity. Before inhibitors identified in this manner can progress into potential antimicrobials, it is necessary to evaluate the defined activity through an *in vitro* screening phase. Primarily, this consists of two general biochemical methodologies: target-based to assess protein-inhibitor interaction and cell-based to screen for whole-cell growth effects [224].

Target-based screening assays are used to verify the inhibitor binds the identified target and has a significant effect on its activity. First, it is important the target can be isolated, characterized and a suitable assay can be developed that is simple and cost-effective while being sensitive enough to detect changes in activity [122]. This is to ensure that the assay is able to screen a large amount of small molecules and the results are reproducible. With our target, *m*-Ddh, purified and enzymatically characterized, we determined a standard enzymatic assay for screening. The reduction of the co-factor NADP⁺ to NADPH is monitored at 340 nm by spectrophotometric analysis and the rate of that reaction determined. Inhibition can then be defined by the corresponding changes in the enzymatic rate. This screening strategy is similar to the HTS performed by GSK for the identification of inhibitors against two enoyl-ACP reductase enzymes, FabI and FabK [126]. By monitoring the consumption of NAD(P)H, an inhibitor displaying slight *in vitro* activity was identified and later optimized for whole-cell activity [225, 226]. Observed target inhibition is used to relate potential pharmacodynamics and

pharmacokinetics of the inhibitor by constructing a dose-response curve. This is used to observe the effect that different concentrations have on the enzyme activity, generating the concentration at which half the enzyme activity is inhibited (IC_{50}) [227]. The IC_{50} allows for the assumption that the inhibition is directly correlated to the inhibitor (non-toxic). Further target-based assays during this stage lead into target-related mechanism of action studies, such as defining the inhibition pattern and binding affinity. Additionally, this data provides the relative inhibitor activity and combined with structural analysis outlines the initial SAR studies [228]. Ultimately, this can be used to understand which parts of the ligand can be altered to improve activity or the pharmacodynamics.

One of the most essential properties of an antibiotic is the ability to cross the membrane and exert a biological effect on the cell. However, inhibitors that show activity against a target may not possess the optimal structural properties for whole cell activity. Whole-cell based assays are used to determine whether a small molecule possess antimicrobial properties. MICs, defined as the lowest concentration that inhibits cell growth, are typically used as a cell-based method to relate inhibition to activity. However, it is important to connect whole-cell activity with a target's inhibition or in other words verify the MoA [229]. This can involve the alteration of the intended target such as over- or underexpression of the target [230], transcriptional analysis of inhibitor treated cells [231] and phenotypic profiling of associated changes in the cell [232]. This can be supported with secondary screenings across multiple species. If screening against a species that lacks the target of interest results in significantly higher MICs, it can be assumed that the small molecule has a specified MoA. This has been successfully applied in drug studies involving FabI/FabL [233], which used over-expression and knockout mutants to discover a novel enoyl-ACP reductase functional protein sensitive to triclosan.

Drug discovery is a multi-step approach. Computational methods can rationally select small-molecules, reducing associated cost and time investments of HTS empirical screening. However, *in vitro* experimental screening cannot be eliminated from the process because it is necessary to identify lead inhibitors for drug development. We previously identified more than 100 small molecules through a CBDD method and purchased 11 of the commercially available, top-ranking compounds for further analysis. In this study, we aimed to determine the *in vitro* activity of the small molecules proposed to target *m*-Ddh in *P. gingivalis*. Applying a two-step biochemical screening approach, we used our standard enzymatic assay to evaluate target-inhibitor interaction followed by whole-cell based assays for the assessment of their antimicrobial activity. Using biochemical screening as a complement to computational approaches provides a comprehensive framework for the initial steps of a drug development approach.

Materials and Methods

Bacterial Strains, plasmids and growth conditions. *P. gingivalis* strain W83, *Prevotella intermedia* strain 17 and *S. sanguinis* strain SK36 were all cultured anaerobically (10% CO₂, 10% H₂, and 80% N₂) at 37 °C in tryptic soy broth (TSB) (Becton Dickinson, Franklin Lakes, NJ) supplemented with 1 µg/ml menadione and 5 µg/ml hemin.

Compounds. The selected small molecules were purchased from Vitas-M Laboratory, Ltd. (Moscow, Russia), Molport (Riga, Latvia) and/or eMolecules (La Jolla, CA, USA), which reported purities over 90%, analyzed by NMR and/or LC- MS. All compounds were re-suspended in DMSO (Sigma–Aldrich, St. Louis, MO, USA) prior to use.

Inhibitor screening and determination of IC₅₀ values. A range of concentrations (0 - 3 mM) of each small molecule inhibitor were added to the standard reaction and the percent of *m*-Ddh enzymatic inhibition was measured by the kinetic assay previously described. Percent inhibition was determined by the formula:
$$\frac{(\text{rate of reaction}_{\text{no inhibitor}}) - (\text{rate of reaction}_{\text{inhibitor}})}{\text{rate of reaction}_{\text{no inhibitor}}} \times 100$$
. The concentration of each inhibitor which caused 50% enzymatic inhibition (IC₅₀) was calculated using PRISM v6.04 software (Graphpad, San Diego, CA) from three independent experiments.

Determination of antimicrobial properties. Minimal Inhibitory Concentration (MIC) assays were performed using a broth microdilution method [234]. *P. gingivalis* or *P. intermedia* cells were grown overnight and the following day diluted 1/10 into fresh medium. *S. sanguinis* cells were grown overnight and the following day diluted 1/100 into fresh medium. Cells were allowed to grow to mid-log phase (OD₆₀₀ ≈ 0.5). Inhibitors were serially diluted in 96-well microtiter plates (Jet Biofil, Genesee Scientific, San Diego, CA) and an aliquot of the cell suspension was added to each well with the inhibitor sample for a final cell count of 1×10⁵ CFU/ml. Plates were incubated either overnight (*S. sanguinis*) or five days (*P. gingivalis* and *P. intermedia*) at 37 °C in anaerobic conditions. The MIC was defined as the lowest concentration of inhibitor that visually reduced cell growth relative to the controls.

Minimal bactericidal concentrations (MBC) were determined by plating bacteria from wells of the MIC assay that showed no visible growth. Samples were plated on tryptic soy agar plates supplemented with 5% sheep blood (Becton, Dickinson, Franklin Lakes, NJ) and incubated at 37 °C in anaerobic conditions for 7 days. MBC was defined as the lowest concentration of inhibitor that resulted in no colony formation/growth.

Time-kill assay. *P. gingivalis* cells were grown overnight and the following day diluted 1/10 into fresh medium. Cells were allowed to grow to mid-log phase ($OD_{600} \approx 0.5$) then diluted to a final cell suspension of 1×10^5 CFU/ml. Inhibitors were added at a concentration of $5 \times$ the MIC determined from the 96-well broth microdilution assay. Samples were taken at different time intervals (0, 0.25, 0.5, 1, 2, 3, 4, 6 and 24 h) and plated on tryptic soy agar plates supplemented with 5% sheep blood (Becton, Dickinson, Franklin Lakes, NJ) using an automated Eddy Jet spiral plater (Neutec Group, Farmingdale, NY). Plates were incubated at 37 °C in anaerobic conditions for 7 days.

SEM analysis of inhibitors exposed *P. gingivalis* cells. Untreated or treated *P. gingivalis* cells were deposited onto a 0.1 μ m disposable Millipore filter to remove medium, and samples were fixed using 2% glutaraldehyde in 0.1 M sodium cacodylate buffer (pH 7.4) for 30 min, followed by 1% osmium tetroxide in 0.1 M sodium cacodylate buffer (pH 7.4). Samples embedded in the filters were then dehydrated in ethanol followed by hexamethyldisilazane (HMDS) and allowed to air dry. The filters were sectioned and mounted onto stubs and coated with gold for three minutes (EMS – 550 Automated Sputter Coater, Electron Microscopy Sciences, Hatfield, PA). Micrographs were taken at 30,000 \times total magnification using a Zeiss EVO 50 XVP scanning electron microscope (Carl Zeiss, Peabody, MA).

Analysis of inhibition mechanisms. Kinetic studies were carried out using the standard kinetic assay for the oxidative deamination of *m*-DAP. Reactions were performed in the absence or presence of inhibitors (0 - 0.4 mM) with varying concentrations of the substrate *m*-DAP or co-substrate NADP⁺. The mode of inhibition and K_i was determined from non-linear regression using PRISM v6.04 software (Graphpad, San Diego, CA) from three independent experiments.

The mode of inhibition was graphically visualized with Lineweaver-Burk or Hanes-Woolf plots according to Cleland kinetics [235].

Results

Evaluation of enzymatic inhibition against *m*-Ddh. The initial screens for the compounds **4** - **14** were performed by individually adding each one to the assay solution. To identify inhibitors with moderate to low activity, the preliminary screening was run at a concentration of 3 mM. Enzymatic activity was measured by the standard assay described in “Materials and Methods” and the % inhibition was calculated in comparison to the untreated enzymatic rate. This resulted in four compounds (**4**, **5**, **6** and **7**) that displayed at least 90% inhibition of enzymatic activity. The other compounds screened displayed 20% or less. These four compounds were then re-screened with a minimum of six concentrations to determine the IC₅₀ value (Table 5). The IC₅₀ values ranged between 100 µM and 1 mM.

It is known that small molecules identified through large structural databases with high IC₅₀ values can be non-specific aggregators, sequestering the enzyme to its surface preventing activity and causing partial denaturation [236, 237]. Aggregation-based inhibition can be reversed through the addition of non-ionic detergents such as Triton X-100. Therefore to assess the selectivity of our enzyme inhibitor, we re-assayed the dose-dependence with the addition of 0.01% Triton X-100, which had no effect on the normal enzymatic activity of *m*-Ddh. The IC₅₀ values for **4**, **5** and **6** showed no significant difference in activity (Table 5).

Table 5. Analysis of *m*-Ddh enzymatic inhibition with active compounds.

	IC₅₀^a	IC₅₀ + Triton X-100^a
Compound 4	156 ± 26 µM	117 ± 13 µM
Compound 5	309 ± 23 µM	399 ± 82 µM
Compound 6	356 ± 16 µM	350 ± 25 µM
Compound 7	1164 ± 174 µM	ND
Compound 8	No inhibition at 3 mM	ND
Compound 9	No inhibition at 3 mM	ND
Compound 10	2% inhibition at 3 mM	ND
Compound 11	13% inhibition at 3 mM	ND
Compound 12	No inhibition at 3 mM	ND
Compound 13	6% inhibition at 3 mM	ND
Compound 14	17% inhibition at 3 mM	ND

ND, not determined

^a Mean of IC₅₀ ± s.d.

n=3

Evaluation of cellular inhibition against *m*-Ddh. To determine if the small-molecule inhibitors displayed antimicrobial activity, we assessed the minimum inhibitory concentration (MIC) using a standard broth microdilution assay. Compounds **4**, **5** and **6** were tested for their ability to visually inhibit growth of *P. gingivalis* cells and Erm was used as a control. Compounds **4** and **5** showed moderate antimicrobial activity with MICs of 250 μ M and 167 μ M, respectively (Table 6). With an MIC over 2 mM, compound **6** was determined not to be appropriate for whole-cell growth inhibition. Testing of the minimum bactericidal concentration (MBC) following the MIC assay, showed only slight differences between the MBC and MIC value (ratio of less than 4:1) for compound **4** and **5**. Compound **6** was not screened for bactericidal activity as the concentration higher than the MIC would have been affected by the solvent concentration.

As a preliminary screen to link growth inhibition to our specific target, we first screened the compounds against *P. intermedia*, another Gram-negative oral pathogen that also possesses the target and *S. sanguinis*, a Gram-positive oral colonizer that does not. Comparison of the data did indicate select growth inhibition. *P. intermedia* showed increased sensitivity to compound **4** but decreased sensitivity to **5**. However, **4** and **5** had the greatest differences between *S. sanguinis* and *P. gingivalis* growth inhibition. Compound **4** showed 7 \times MIC of *P. gingivalis* and **5** was more than double. Compound **6** didn't show significant levels of inhibition similar to *P. gingivalis* and *P. intermedia* (Table 6). Previous studies show that supplementation of *m*-DAP into to the media of mutants lacking one or more of the enzyme within the *m*-DAP/lysine biosynthesis pathway can allow for growth and recovery [180, 183, 238]. Thus, to further link growth inhibition with *m*-Ddh inhibition, we re-assessed the MICs with the addition of the target's essential product *m*-DAP. This had no effect on the growth inhibition for **5** or **6** in *P. gingivalis*, *P. intermedia* or *S. sanguinis*. However, compound **4** showed a slight increase in

Table 6. Analysis of whole-cell inhibition with active inhibitors

<i>P. gingivalis</i>						<i>P. intermedia</i>				<i>S. sanguinis</i>		
0 μ M <i>m</i> -DAP		25 μ M <i>m</i> -DAP	200 μ M <i>m</i> -DAP	0 μ M <i>m</i> -DAP	25 μ M <i>m</i> -DAP	200 μ M <i>m</i> -DAP	0 μ M <i>m</i> -DAP	25 μ M <i>m</i> -DAP	200 μ M <i>m</i> -DAP	0 μ M <i>m</i> -DAP	25 μ M <i>m</i> -DAP	200 μ M <i>m</i> -DAP
Compound 4	250	8	34	8	2	1080	1080	1080	1080	1080	1080	1080
Compound 5	167	252	252	505	505	505	505	505	505	1010	1010	1010
Compound 6	>2000	>2000	1080	1080	1080	1080	>2000	>2000	>2000	>2000	>2000	1080
Erm	<0.244 μ g/ml	0.488 μ g/ml	<0.244 μ g/ml									

<i>P. gingivalis</i>	
MBC (μ M)	
Compound 4	1740
Compound 5	305
Compound 6	ND
ND, not determined n=3 or 6	

sensitivity when 25 μ M of *m*-DAP was added to the assay for *P. gingivalis* and *P. intermedia* (Table 6).

As MICs are static measurements, we wanted to next observe the kinetic effects the three active inhibitors had on the growth of *P. gingivalis*. Cells were exposed to the compounds and total viable cells were measured at different time intervals. At 5 \times MIC, compound **4** reduced the viable *P. gingivalis* cell count by 2 log₁₀ CFU /ml within 6 h of exposure. However, **5** rapidly reduced the cell count upon treatment. After 2 h of exposure there was a 5 log₁₀ CFU/ml reduction, resulting in no viable cell count (Fig. 17). Cells exposed to **6** at the higher concentration treatments were affected by the DMSO solvent and could not be assessed.

Phenotypic profiling linking target inhibition to growth effects were assessed by changes in the cellular morphology. *P. gingivalis* cells exposed to either compound **4** or **5** were examined by scanning electron micrograph (SEM). Compound **6** was left out of the analysis due to the high MIC concentration. Treated *P. gingivalis* showed an alteration of the cellular structure compared to the untreated cells (Fig. 18). Cells were visibly misshapen.

SAR evaluation of compound 4 core scaffold. From our target-based screening assay, compound **4** displayed the most potent activity. Therefore, commercially available analogues with 70% -90% structural similarities identified through the ZINC database were purchased and screened for a preliminary evaluation of the core scaffold SAR. The structures and activities are listed in Table 7. The analogues differ at two substitution sites at both ends of the structure with IC₅₀ ranging from 127 – 238 μ M. Removing the methoxy group from the R₁ phenyl group of the

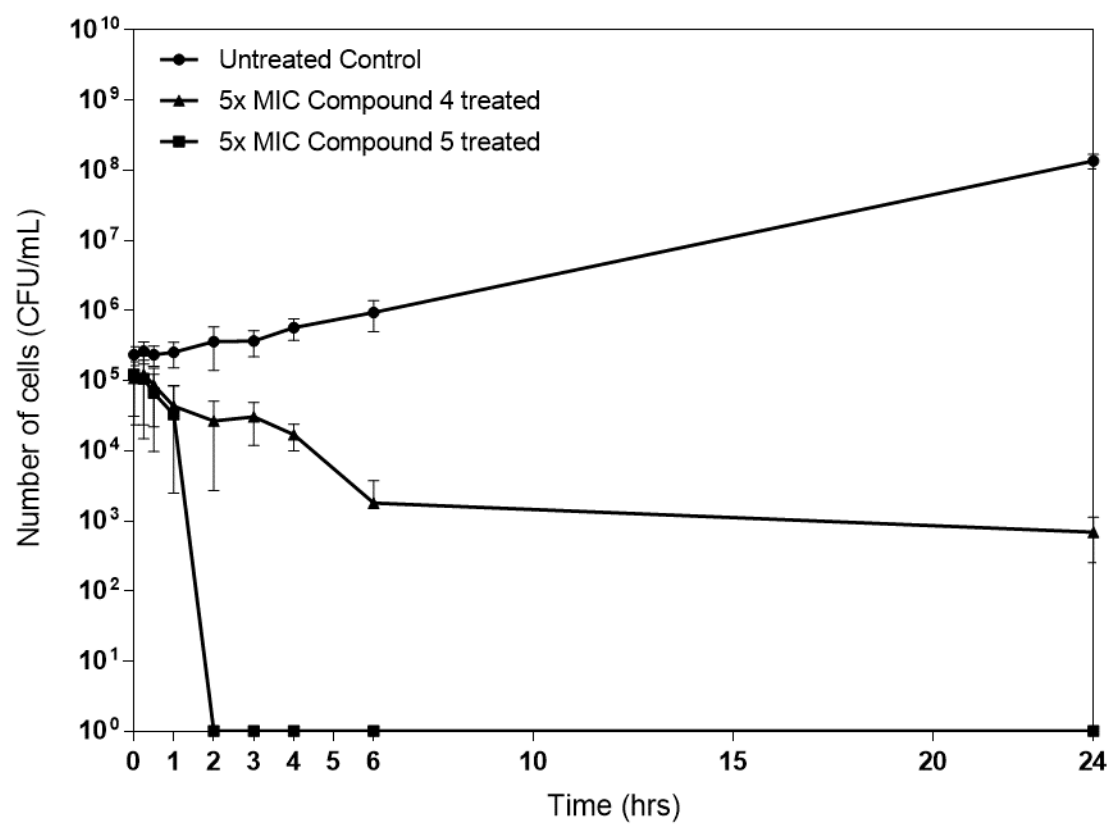
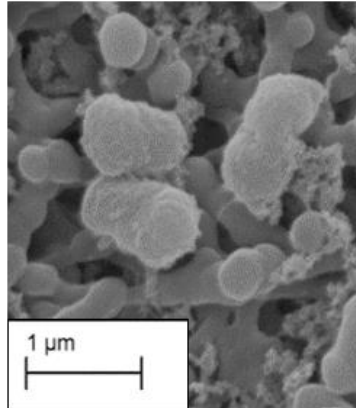


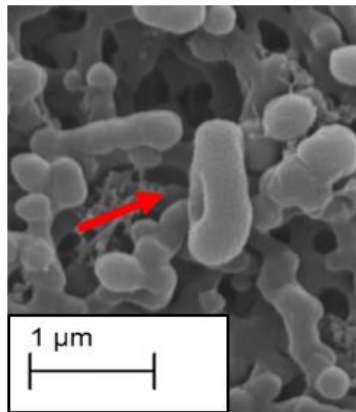
Figure 17. Time-kill analysis of compound treated *P. gingivalis*.

P. gingivalis cells were treated with 5x the previously determined MIC for either Compound **4** (triangle) or Compound **5** (square) and bacterial cell counts were assessed at 0, 0.25, 0.5, 1, 2, 3, 4, 6 and 24 hours. The mean plus the standard deviation is shown for each time point from a minimum of n=3 independent experiments. For cell counts equal to 0 CFU/mL, 1 was used for the log transformation.

a



b



c

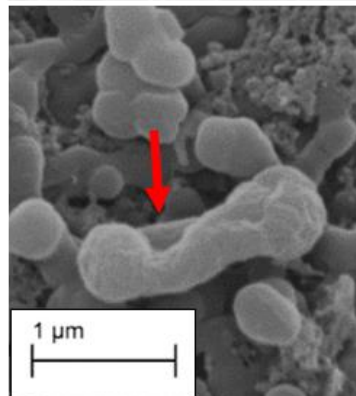
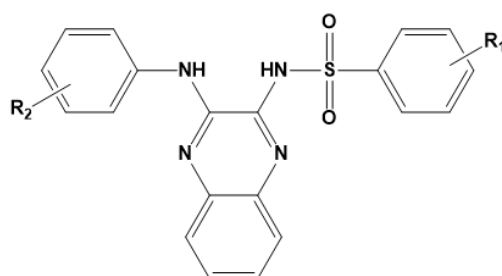


Figure 18. SEM analysis of compound treated *P. gingivalis* cells.

(a) Untreated cells. **(b)** Compound **4** treated cells at 5x the previously determined MIC concentration. **(c)** Compound **5** treated cells at 5x the previously determined MIC concentration.

Table 7. SAR of compound 4 analogues.



	R₁	R₂	HINT	IC₅₀ (μM)^a
Compound 4a	4-methyl	3-carboxyl	3397	127 ± 9
Compound 4b	4-methyl	4-carboxyl	3402	143 ± 6
Compound 4c	H	3-hydroxy-4-carboxyl	2774	155 ± 4
Compound 4d	4-methyl	3-hydroxy-4-carboxyl	3006	158 ± 7
Compound 4e	3,4-dimethyl	3-carboxyl	3280	237 ± 18

^a Average IC₅₀ ± s.d.
n=3

original structure and replacing it with a methyl group resulted in the most potent inhibitor screened with a slight decrease in the IC_{50} (compound **4a**, IC_{50} = 127 μ M). This may be due to a decrease in size preventing a steric hinderance. However, adding electron withdrawing 3- or 4-carboxyl or electron donating 3-hydroxyl groups on the core R_2 phenyl moiety, didn't exhibit any significant effect on the activity. This was observed from the structures of compound **4a-4d** suggesting that substitutions are tolerated around the ring within the binding site of these compounds. However, the addition of methyl groups to the 3- and 4- position on the R_1 phenyl moiety for **4e** resulted in an increase in the IC_{50} (IC_{50} = 238 μ M). Comparison with between compound **4a** and **4e** suggest that an additional moiety to the 3-position of the phenyl ring is less tolerable to the intermolecular interaction.

Evaluation of inhibition mechanism. Based on our pharmacophore model and docking studies, our three active inhibitors share a similar binding pattern within the active site, competing with *m*-DAP. The intermolecular interactions are depicted in Figure 19. Hydrogen bond interactions are represented by dashed lines. The interactions are shown to follow the proposed pharmacophore models in that hydrogen bonding interactions between the sulfonamides are occurring with Arg183 and Thr173. Hydrophobic stacking interactions occur between the aromatic rings of the inhibitors and residues Phe148 and potential Trp123 of the active site. There may also be potential hydrogen bonding between the ligands' carboxylic groups and residues Gly155 and Met154. In addition, for our model, the carboxylate groups form hydrogen bonds with the backbone amide of Ser153 and Met154. Compound **4** makes an additional hydrogen- bonding interaction with Tyr207 and π - π stacking interaction with Phe148, which

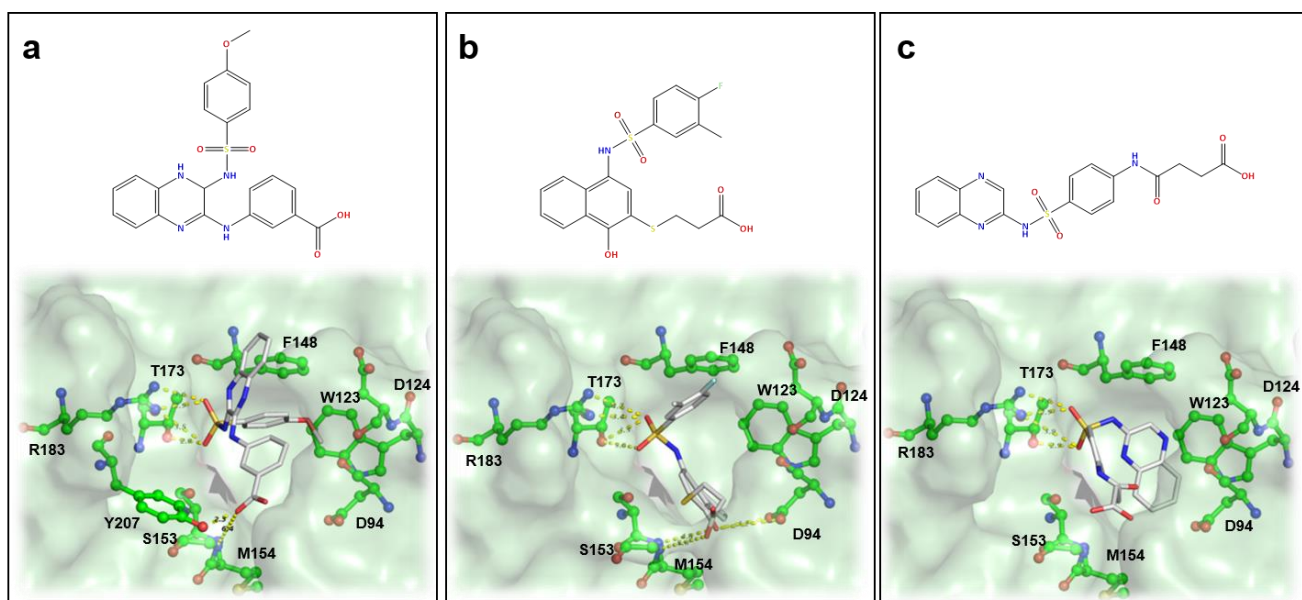


Figure 19. 3D analysis of binding model for active inhibitors.

Optimal binding model for Compound **4** (a), Compound **5** (b) and Compound **6** (c) are shown above. Key residues are displayed as ball and sticks and colored corresponding to atom type. Residues labeled in the bottom middle figure correspond to all figures. Hydrogens were omitted for clarity. Potential hydrogen bonding interactions between *m*-Ddh residues and inhibitors are shown by yellow dashed lines.

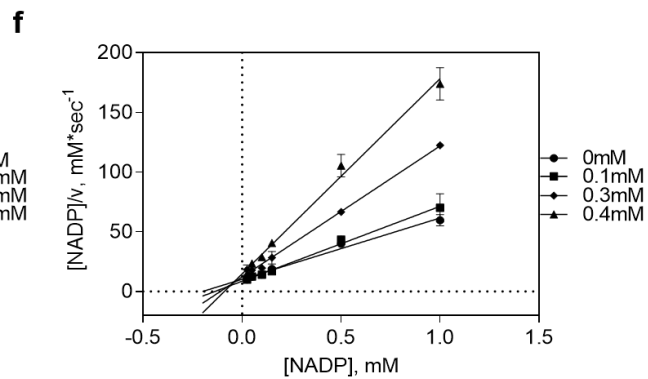
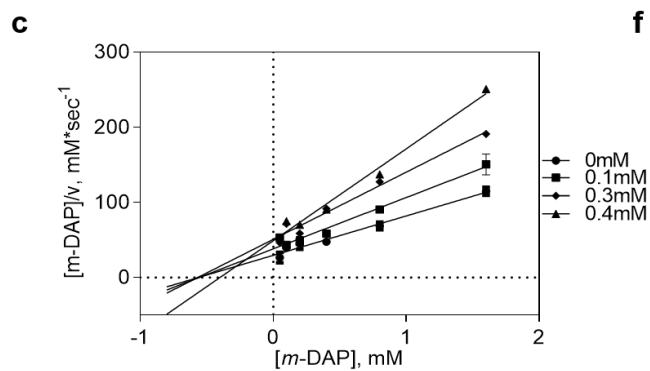
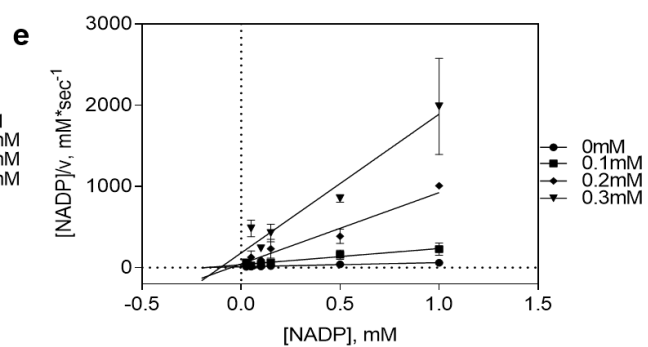
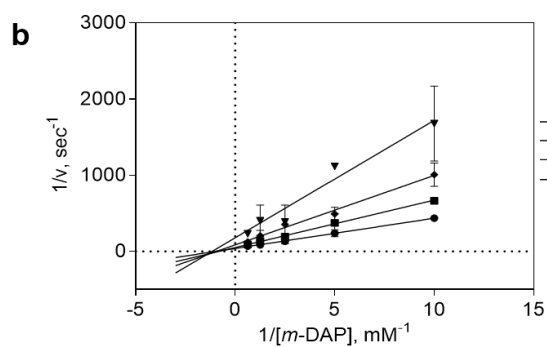
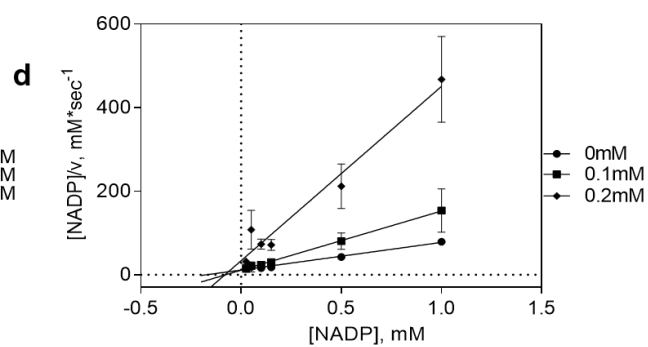
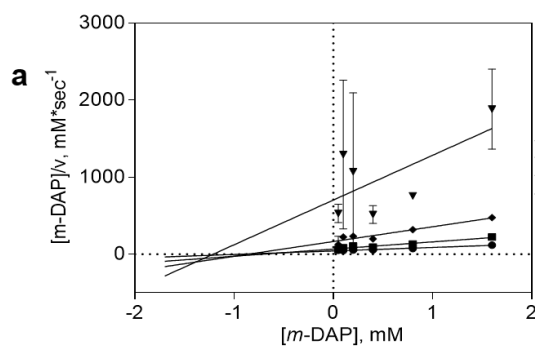


Figure 20. Inhibition mechanism of active compounds in regards to substrate, *m*-DAP and co-substrate, NADP⁺.

(a) Compound **4** **(b)** Compound **5** and **(c)** Compound **6** inhibition mechanisms against *m*-DAP. **(d)** Compound **4** **(e)** Compound **5** and **(f)** Compound **6** inhibition mechanisms against NADP⁺.

may be one of the many reasons for its better activity. To verify this interaction, the method of inhibition was kinetically determined. The substrate *m*-DAP was varied over several concentrations of each inhibitor. The inhibition pattern for all three compounds was revealed to be non-competitive, as there was no change in the K_m and a decrease in V_{max} compared to the activity in the absence of the inhibitor (Fig. 20a-c). Due to structural similarities between the compound scaffolds and the co-substrate and the proximity of the two binding sites, we next examined the possibility that one of more of the inhibitors may be binding in the NADP⁺ binding site. This inhibition study showed an uncompetitive inhibition pattern as there was an increase in the K_m and a decrease in V_{max} compared to the enzymatic rate in the absence of the inhibitor (Fig. 20d-f).

Discussion

In this study, we evaluated the activity of compounds identified through a CBDD method. A two-step screening strategy was employed to assay *m*-Ddh specific inhibition and *P. gingivalis* growth inhibition. From 11 compounds purchased and screened *in vitro*, three showed target-specific inhibition with IC₅₀ values ranging from 100-300 μ M. Several other compounds showed slight activity in the higher millimolar range. The most potent compounds (**4**, **5** and **6**) showed limited structural similarity. The core structure of compound **4** and **6** were both sulfonyl amino quinoxaline derivatives while **5** was a sulfonyl amino naphthalene derivative. However, all three possessed similar functional groups and were predicted to share basic pharmacophoric features. The three compounds possessed a sulfonamide core attached to large aromatic structures with carboxylate functional groups. The importance of the sulfonamide has literature precedence in the search for antimicrobials targeting lysine biosynthesis. Compounds structurally similar to our

hits that possessed sulfonamides and sulfones were identified as fairly good inhibitors of dihydrodipicolinate reductase, another enzyme in the lysine biosynthesis pathway [239, 240]. Based on docking studies, the sulfonamide groups were predicted to favorably interact with Arg183 and Thr173 forming hydrogen bonds, while the aromatic moieties would create hydrophobic interactions. From docking studies it appears the distance of the carboxylic groups may not allow for hydrogen bonds with residues at the other end of the binding pocket. However, based on other studies, *m*-Ddh exists in an open and closed conformation [218]. This would bring the residues surrounding the compounds within the distance necessary to form stronger hydrogen bonds. It should be noted that previous studies have reported more potent inhibitors against *m*-Ddh [209, 239, 241]. However, these compounds are typically small analogous structures, derived from the substrate *m*-DAP that possess few ‘drug-like’ features, making optimization difficult [239], and suggesting little hope for selectivity. Our compounds allow for the development of more active compounds, similar to the *in silico* screening against thymidylate synthase, an enzyme is essential for DNA replication, by DesJarlais *et al.* [242]. The initial computational study yielded several compounds with activity in the high micromolar range, but following further analysis and optimization resulted in an increase in potency as well as verification of the binding mode.

In the pursuit of a preliminary SAR analysis, 5 analogues of the most active inhibitor (compound **4**) scaffolds were screened. Through a small substitution, we identified one compound (**4a**) that showed a slight increase in activity and one compound with a decrease in activity (**4e**). Based on our predicted knowledge of the compound binding, comparison of the intermolecular interaction of **4a** and **4e** indicate, while the binding model appears similar, experimentally it appears the additional functional groups around the phenyl ring closer to

Trp123 and Phe148 within the binding pocket are not favored. On the other end, additional groups to the phenyl ring near Thr173 and Gly155 are tolerated. Screening of an analogue with a 3-methyl addition at the R₁ substitution and 3-carboxyl at the R₂ substitution would allow for the direct comparison to **4a**. This would allow us to determine if the methyl group in the para-position increases the potency and if the additional methyl group decreases potency. The addition of a hydroxyl group or a carboxylate in the 3- or 4-position showed no significant difference in activity. As this is a preliminary SAR study, screening analogues with increased structural diversity or fragment-based screening based on the deconstruction of the core scaffold will provide a more comprehensive analysis of the essential interactions.

Previous initial-velocity data [216] in *m*-Ddh show that the reaction proceeds through a sequential ordered ternary-binary mechanism with NADP⁺ binding first, followed by the substrate *m*-DAP. The product is then released, followed by NADPH. Our studies into the mechanism of inhibition show the inhibitory compounds to be non-competitive with respect to *m*-DAP, but uncompetitive with respect to NADP⁺. In concordance with the binding order, this would indicate that the compounds bind to either the Enzyme-NADP⁺ complex or the Enzyme-NADPH complex, thus potentially preventing a necessary conformational change and/or reducing the affinity of *m*-DAP for the protein. This type of mechanism of inhibition could be beneficial for future therapeutics. Treatment with an optimized inhibitor competing with *m*-DAP, would result in the accumulation of the substrate within the cytosol or a regulation mechanism in order to meet the need of the cell. An increase in the localized substrate would then need to be balanced by high concentrations of the inhibitor. A non-competitive inhibitor, however would not be affected by the increased concentration of substrate compared to a competitive inhibitor, making it more effective at lower concentrations.

One of the most difficult aspects of target-based drug discovery is identifying compounds that show effective whole-cell activity while maintaining the key pharmacokinetics. While the compounds identified showed potential against *m*-Ddh, they exhibited only slight antimicrobial activity in *P. gingivalis*. Nevertheless, the potential for antimicrobial activity should not be ignored as analogous structures and optimization of the scaffold could improve whole cell inhibition. Several reasons could contribute to the low observed MIC values. For one, bacterial inhibitors must be able to penetrate the cell membrane while maintaining enough soluble and free fractions to inhibit the target at sufficient concentrations. The compound also must avoid being expelled from the cell through efflux pumps. Another reason could be due to the sulfonamide group present on compounds **4**, **5** and **6**. Sulfonamides are well known antimicrobials that target folate biosynthesis, and bacterial cells may show a degree of drug resistance [239]. There is also the potential for non-specific inhibition or off-target interactions. This would result in what appears to be a variation in activity between whole-cell and target inhibition. This was observed for compound **5** which displayed a lower MIC than IC₅₀. While a detailed structure-activity relationship for the antimicrobial properties cannot yet be determined from these studies, it may be speculated that the more favorable whole-cell activity seen in Compound **5** compared to compound **4** and **6** is due to lipophilicity. The relatively low lipophilic nature of compound **6** (cLogP = 1.70) compared to **5** (cLogP = 3.68) may have decreased permeability through the cellular membrane of *P. gingivalis*. While compound **4** displayed the more potent target-based screening, the high lipophilic nature (cLogP = 5.26) may have had a significant effect on the solubility, reducing the efficacy during cell-based screening [243]. However, we were able to show differential activity indicating specificity. Testing in *P. gingivalis* and *P. intermedia*, both of which possess the target, *m*-Dh showed a greater degree of

growth inhibition compared to *S. sanguinis*, which lacks the target. Compounds tested for *S. sanguinis* showed almost no antimicrobial activity (compound **4** and **6**) or MICs more than double that of those seen in *P. gingivalis* (compound **5**). Supplementing *m*-DAP into the growth media as a substrate competitor to the compounds did not correlate to an increase in MICs. While previous studies show that supplementation of *m*-DAP for *E. coli* and *M. smegatis* mutants restores growth in DAP auxotrophs [180, 238], it is not known whether *P. gingivalis* can specifically uptake and utilize *m*-DAP for its benefit. In addition if the compounds act in a non-competitive or irreversible manner, addition of *m*-DAP would not necessarily compensate for the inhibitor activity. Analysis of the kinetic growth inhibition indicated the activity of these compounds are time and dose dependent, with higher concentrations and longer exposure times leading to an increased loss of cell viability. Compound **5** completely eliminated *P. gingivalis* cell viability after two hours of exposure at 5x the MIC concentration, while Compound **4** maintained a low cell count after six hours of treatment. This is also consistent with the MBC being less than 4x the MIC as antimicrobials with MBC in close range of the MIC are typically classified as bactericidal.

In conclusion, we demonstrate the rationale for applying a focused biochemical screening with CBDD to identify inhibitors. Our results show that by utilizing a target screen we can select compounds which actively inhibit *m*-Ddh from *P. gingivalis*, identify the inhibition mechanisms and assess the SAR. Coupled with whole-cell screen we can identify compounds with antimicrobial properties and link the MoA. While the activities of these first generation compounds are somewhat weak, continued studies into the intermolecular binding interaction could help to discern which features are key, allowing for the improvement of novel inhibitors.

Conclusion

The goal of this project was to introduce the beginnings of a model for antimicrobial drug discovery using the periodontal pathogen, *P. gingivalis*. Accordingly, we have successfully presented a thorough strategy. By understanding the three main biological functions for bacterial survival, we were able to create a network of pathways connected to the synthesis of essential metabolic components. As we demonstrated, this allowed for the accurate prediction of essential genes as potential antibacterial targets. Additionally, we believe this protocol allows for the comparison of alternative pathways and gene sets for species-selective targeting. Following the selection of a target, rational computer-based drug discovery (CBDD) was applied for the rapid and cost-effective identification of small-molecule compounds. We continued with our model to show that analysis of the protein structure and intermolecular interactions allowed for the utilization of a high-throughput virtual screen to select compounds with a higher probability of demonstrating target inhibition. Finally, by developing a simple and effective screening strategy to assess target and cell growth inhibition, we established that we could determine the *in vitro* activity of selected small-molecule compounds. Through this screening, not only are compounds with activity verified, but the mechanism of inhibition is determined and SAR studies can be started.

The significance of our study should also be noted for our chosen model. First, we identified *m*-Ddh, an essential enzyme for *P. gingivalis*, as a potential pathogen-specific target within the oral cavity. The dual role of *m*-Ddh in protein and cell wall biosynthesis makes it an exceptional drug target, as it would inhibit two essential biological processes. Additionally, *m*-Ddh is present within several other pathogenic colonizers, including *P. intermedia*, *T. denticola* and *T. forsythia*, while absent in the majority of early oral colonizers. Since *m*-Ddh is a unique

target found within a limited number of species, we theorize this would allow us to potentially develop novel narrow-spectrum therapy for the treatment of periodontal disease by specifically targeting key pathogens yet preserving the healthy microbiome. Second, to the best of our knowledge, this is the first computationally motivated target-based drug discovery for this periodontal pathogen. CBBD is a successful strategy, especially in systemic diseases [202]. By applying CBBD we were able to identify over 100 small molecule compounds with the potential of target-specific inhibition. Screening of the top-ranked, commercially available compounds, resulted in a 36% hit rate. Demonstrating that a CBDD can be applied to this bacterium, we hope this will drive new focus into drug therapies for oral diseases utilizing a rational approach. Last, *P. gingivalis* and periodontal disease can be used as the starting model for rational species selective drug discovery. As the human body is colonized by billions of microbial cells, it is highly unlikely a bacterial infection, especially a biofilm, would be isolated into a single species. Therefore, careful consideration of the effect therapeutic treatment has on the microbiome should be taken, bringing into play the justification for targeted therapies. We theorized periodontal disease and *P. gingivalis* would be a beneficial model as the oral microbiome is one of the most diverse sites on the body and *P. gingivalis* is widely recognized as one of the major contributors to periodontitis [32, 34, 70, 86, 244]. *P. gingivalis* can easily be cultured in the lab within a reasonable amount of time, allowing for *in vitro* testing. Additionally, periodontal disease has been extensively studied for years yielding suitable models for the disease, both *in vitro* and *in vivo* [245-247].

Some challenges encountered during this study should be addressed as they may add in the future improvement of the process. For one, it is important to verify that the target is essential *in vivo* and *in vitro*. Essential gene studies are usually preformed under laboratory,

nutrient-rich conditions. This does not always correspond to *in vivo* conditions where nutrients may be provided by the host or alternate pathways induced through stress factors. This is especially important for broad-spectrum targeting across multiple species. If possible, *ex vivo* models can be established that more closely mimic *in vivo* growth conditions and gene essentiality can be observed in various nutrient conditions. Subsequently, rational drug discovery assays must be designed to verify target inhibition during cell-based screening. The use of WT and mutant bacteria strains that overexpress and underexpress the target will not only identify compounds with antibacterial properties but select for those with target-specific inhibition. Alternatively, inducible antisense RNA could be used to control the expression level of the target and sensitivity to different inhibitor concentrations can be gauged. The compound database screened also needs careful consideration. Through studies by Lipinski *et al.*, it was observed that these inhibitors generally follow certain traits (no more than 500 Da in size, five or fewer hydrogen bond donors, no more than 10 hydrogen bond acceptors and a logP_{o/w} no greater than five), known as Lipinski's rule of five [187, 188]. However, antimicrobials do not always follow this rule [97, 126]. Analysis of clinically approved antibiotics show that they are structurally more complex and larger in size. They also possess more hydrogen donors and acceptors and are more hydrophilic [97, 126]. Increased structural diversity in established chemical libraries or the construction of dedicated databases with compounds possessing antimicrobial properties would be beneficial. There has also been a resurgence in the search for natural compounds, similar to those identified during the 'golden era' of antibiotic research [97, 248]. Combining these compounds with a CBDD approach and other technical advances may bring new interest and success in this field of research.

We anticipate several directions for future studies. We first intend to continue the evaluation of the MoA to link whole-cell activity to target inhibition. By placing *m*-Ddh under an inducible promoter we can regulate the expression and screen compounds for changes in whole-cell growth inhibition. A more focused SAR study is also underway through the deconstruction/reconstruction of the three active compound inhibitors. The scaffold of the three active compounds will be systematically broken down into core functional groups. The basic structure of these compounds will then be re-evaluated through *in silico* and *in vitro* screening to assess the differences in activity and potency. The compound will then be reconstructed to include only the essential functional groups. Following an increase in target-based inhibition, the structure can be further optimized for increased whole-cell activity. We also plan to further evaluate the protein-compound binding. Select mutants will be made near the binding pocket of *m*-Ddh that do not have a significant impact on the catalytic efficiency. Compounds will then be re-screened against the mutants and inhibition will be measured to determine if the mutant affected compound binding compared to WT. We also screened a relatively small number of compounds from the results of our SBVS. Therefore, we plan to have a second round of screening from the remaining hits to increase our pool of active compound inhibitors.

To conclude, we suggest that this method is an interdisciplinary approach whose application has the potential to extend beyond what has been shown here. Employing a combination of microbiologists, medicinal chemists and bioinformaticians we believe with time this approach can flourish. Using their knowledge in the necessary fields and taking the backbone of this project, we hope this can be applied to increasing deadly pathogens, such as *M. tuberculosis*, as an alternative method in antibacterial research.

Appendix

Supplemental Table 1. Predicted putative essential genes in *P. gingivalis* strain W83.

Functional category	Pathway	KEGG Pathway	KEGG KO#	GeneID	Name	Protein ID
Cell envelope (cell membrane)	Peptidoglycan biosynthesis	Peptidoglycan biosynthesis	K00075	PG1342	murB	UDP-N-acetylenolpyruvoylglucosamine reductase E
			K01924	PG0581	murC	UDP-N-acetylmuramate--L-alanine ligase E
			K01925	PG0578	murD	UDP-N-acetylmuramoyl-L-alanyl-D-glutamate synthetase E
			K01928	PG0576	murE	Mur ligase family protein E
			K03340	PG0806	-	Gfo/Idh/MocA family oxidoreductase E
			K01921	PG0729	ddl	D-alanyl-alanine synthetase A E
			K01929	PG1106	murF	D-Ala-D-Ala adding enzyme E
			K01000	PG0577	mraY	phospho-N-acetylmuramoyl-pentapeptide-transferase E
			K02563	PG0580	murG	N-acetylglucosaminyl transferase E
			K05366	PG0794	pbp1a	penicillin-binding protein 1A UA
			K03587	PG0575	-	penicillin-binding protein 2 E
			K00215	PG2002	dapB	dihydrodipicolinate reductase E
			K00928	PG2189	lysC	aspartate kinase E
			K00133	PG0571	asd	aspartate-semialdehyde dehydrogenase E
			K01714	PG2052	dapA	dihydrodipicolinate synthase E
		D-Glutamine and D-glutamate metabolism	K01776	PG0705	murI	glutamate racemase E
		D-Alanine metabolism	K01775 K01929	PG1097	alr	putative bifunctional UDP-N-acetylmuramoyl-tripeptide:D-alanyl-D-

						alanine ligase/alanine racemase	
		Terpenoid backbone biosynthesis (MEP/DOXP pathway)	K01662	PG2217	dxs	1-deoxy-D-xylulose-5-phosphate synthase	E
			K00099	PG1364	dxr	1-deoxy-D-xylulose 5-phosphate reductoisomerase	E
			K00991	PG1434	ispD	2-C-methyl-D-erythritol 4-phosphate cytidyltransferase	E
			K00919	PG0935	ispE	4-diphosphocytidyl-2-C-methyl-D-erythritol kinase	E
			K01770	PG0028	ispF	2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase	E
			K03526	PG0952	ispG	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase	E
			K03527	PG0604	ispH	4-hydroxy-3-methylbut-2-enyl diphosphate reductase	UA
		Terpenoid backbone biosynthesis	K00806	PG0190	uppS	undecaprenyl pyrophosphate synthetase	E
		Polysaccharide transporter	-	PG0117	-	polysaccharide transport protein	E
	Lipopolysaccharide biosynthesis	Lipopolysaccharide biosynthesis	K00677	PG0070	lpxA	UDP-N-acetylglucosamine acyltransferase	E
			K02372	PG0071	lpxC	bifunctional UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosamine deacetylase/(3R)-hydroxymyristoyl-ACP dehydratase	E
			K02536	PG0072	lpxD	UDP-3-O-[3-hydroxymyristoyl] glucosamine N-acyltransferase	E
			K00912	PG0638	lpxK	tetraacyldisaccharide 4'-kinase	E
			K02527	PG1565	-	3-deoxy-D-manno-octulosonic-acid transferase	E
			K02517	PG2222		acyltransferase	E
	Fatty acid biosynthesis	Pantothenate and CoA biosynthesis	K00954	PG0369	coaD	phosphopantetheine adenylyltransferase	E
			K00859	PG0483	coaE	dephospho-CoA kinase	E

Energy production	Glycerophospholipid / glycerolipid metabolism	Biotin metabolism	K03524	PG1601	birA	biotin--acetyl-CoA-carboxylase ligase	E
		Fatty acid biosynthesis	K02078	PG1765	acpP	acyl carrier protein	E
			K00648	PG2141	fabH	3-oxoacyl-(acyl carrier protein) synthase III	E
			K00645	PG0138	fabD	malonyl-CoA:ACP transacylase,	E
			K09458	PG1764	-	3-oxoacyl-(acyl carrier protein) synthase II	E
			K00059	PG1239	fabG	3-ketoacyl-(acyl-carrier-protein) reductase	UA
			K02371	PG1416	fabK	enoyl-acyl carrier protein(ACP) reductase,	UA
		Glycerophospholipid metabolism	K01803	PG0623	tpiA	triosephosphate isomerase	E
			K00057	PG1369	gpsA	glycerol-3-phosphate dehydrogenase	E
			K00980	PG2068	tagD	glycerol-3-phosphate cytidyltransferase	E
			K00655	PG1249	plsC	1-acyl-sn-glycerol-3-phosphate acyltransferase,	UA
	Glycolysis	Glycerophospholipid metabolism	K00981	PG0046	cdsA	phosphatidate cytidyltransferase,	E
		Glycolysis / Gluconeogenesis	K01810	PG1368	pgi	glucose-6-phosphate isomerase	E
			K01835	PG2010	pgm	phosphomannomutase	UA
			K04041	PG0793	fbp	fructose-1,6-bisphosphatase	E
			K01624	PG1755	fba	fructose-bisphosphate aldolase	E
			K00134	PG2124	gapA	glyceraldehyde 3-phosphate dehydrogenase	E
			K00927	PG1677	pgk	phosphoglycerate kinase	E
			K01834	PG0130	gpmA	2,3-bisphosphoglycerate-dependent phosphoglycerate mutase	E
			K15634	PG1513	-	phosphoribosyltransferase/phosphoglycerate mutase	E
			K01689	PG1824	eno	phosphopyruvate hydratase	UA
			K01006	PG1017	ppdk	pyruvate phosphate dikinase	NE

			K01610	PG1676	pckA	phosphoenolpyruvate carboxykinase	NE
	Pentose phosphate pathway	Pentose phosphate pathway	K00948	PG2097	prsA	ribose-phosphate pyrophosphokinase	E
			K01619	PG1996	deoC	deoxyribose-phosphate aldolase	UA
	NAD ⁺ /NADP ⁺ biosynthesis	Nicotinate and nicotinamide metabolism	K00763	PG0057	pncB	nicotinate phosphoribosyltransferase	E
			K01950	PG0531	nadE	NAD synthetase	E
			K00858	PG0629	ppnK	inorganic polyphosphate/ATP-NAD kinase	E
	Genetic Information inheritance	Purine metabolism	K00942	PG0512	gmk	guanylate kinase	E
			K00939	PG0791	adk	adenylate kinase	E
		Pyrimidine metabolism	K09903	PG1902	pyrH	uridylate kinase	E
			K00384	PG1134	trxB	thioredoxin reductase	UA
			K00560	PG2060	thyA	thymidylate synthase	E
		Purine metabolism/Pyrimidine metabolism	K00945	PG0603	cmk	cytidylate kinase	E
			K00525	PG1129	nrd	ribonucleotide reductase	NE
		Folate biosynthesis	K01495	PG0625	folE	GTP cyclohydrolase I	E
			K00950	PG1541	folK	2-amino-4-hydroxy-6-hydroxymethyldihydropteridine pyrophosphokinase	UA
			K01633	PG2091	folB	dihydroneopterin aldolase	E
			K00796	PG1589	folP	dihydropteroate synthase	E
			K01930	PG0463	folC	dihydrofolate synthetase	UA
			K00287	PG2061	folA	dihydrofolate reductase	E
		DNA replication (DNA polymerase III)	K02343	PG1418	dnaX	DNA polymerase III subunits gamma and tau	E
			K02341	PG0932	-	DNA polymerase III subunit delta'	UA
			K02340	PG0949	-	DNA polymerase III subunit delta	E
			K02342	PG1852	-	exonuclease (DNA polymerase III subunit epsilon)	E
			K02342	PG0223	-	exonuclease (DNA polymerase III subunit epsilon)	NE
			K02337	PG0035	dnaE	DNA polymerase III DnaE	UA
			K02338	PG1853	dnaN	DNA polymerase III subunit beta	E

			K02314	PG1242	dnaB	replicative DNA helicase	E
			K02316	PG1814	dnaG	DNA primase	E
			K03111	PG0271	ssb	single-strand DNA-binding protein	E
			K01972	PG1253	ligA	NAD-dependent DNA ligase LigA	E
			K02335	PG1794	polA	DNA polymerase I	E
		DNA replication proteins (Initiation and re-initiation)	K02313	PG0001	dnaA	chromosomal replication initiation protein	E
			K03530	PG0121	hup-1	DNA-binding protein HU,	UA
			K03530	PG1258	hup-2	DNA-binding protein HU	E
		DNA replication proteins (DNA topoisomerase)	K02470	PG1702	gyrB	DNA gyrase subunit B	E
			K02469	PG1386	gyrA	DNA gyrase subunit A	E
			K03168	PG0754	topA	DNA topoisomerase I	E
			K02622	PG0368	parE	DNA topoisomerase IV subunit B	E
			K02621	PG1622	parC	DNA topoisomerase IV subunit A	E
		Homologous recombination (Holliday junction)	K03551	PG0488	ruvB	Holliday junction DNA helicase B	E
			K07447	PG2202	-	Holliday junction resolvase-like protein	UA
	Transcription	RNA polymerase	K03043	PG0394	rpoB	DNA-directed RNA polymerase subunit beta	UA
			K03046	PG0395	rpoC	DNA-directed RNA polymerase subunit beta'	E
			K03086	PG0594	rpoD	RNA polymerase sigma factor RpoD	UA
			K03040	PG1911	rpoA	DNA-directed RNA polymerase subunit alpha	E
	Protein biosynthesis	Ribosome	K02863	PG0391	rplA	50S ribosomal protein L1	E
			K02886	PG1935	rplB	50S ribosomal protein L2	E
			K02906	PG1938	rplC	50S ribosomal protein L3	E
			K02926	PG1937	rplD	50S ribosomal protein L4	E
			K02931	PG1926	rplE	50S ribosomal protein L5	E
			K02933	PG1923	rplF	50S ribosomal protein L6	E
			K02935	PG0393	rplL	50S ribosomal protein L7/L12	E
			K02864	PG0392	rplJ	50S ribosomal protein L10	E

			K02867	PG0390	rplK	50S ribosomal protein L11	E
			K02871	PG0375	rplM	50S ribosomal protein L13	E
			K02874	PG1928	rplN	50S ribosomal protein L14	E
			K02878	PG1931	rplP	50S ribosomal protein L16	E
			K02879	PG1910	rplQ	50S ribosomal protein L17	UA
			K02881	PG1922	rplR	50S ribosomal protein L18	E
			K02884	PG0037	rplS	50S ribosomal protein L19	UA
			K02887	PG0989	rplT	50S ribosomal protein L20	E
			K02890	PG1933	rplV	50S ribosomal protein L22	E
			K02892	PG1936	rplW	50S ribosomal protein L23	E
			K02899	PG0315	rpmA	50S ribosomal protein L27	E
			K02914	PG0656	rpmH	50S ribosomal protein L34	UA
			K02919	PG1915	rpmJ	50S ribosomal protein L36	E
			K02916	PG0990	rpmI	50S ribosomal protein L35	E
			K02895	PG1927	rplX	50S ribosomal protein L24	E
			K02897	PG0167	rplY	50S ribosomal protein L25	E
			K02876	PG1919	rplO	50S ribosomal protein L15	E
			K02907	PG1920	rpmD	50S ribosomal protein L30	E
			K02888	PG0314	rplU	50S ribosomal protein L21	E
			K02904	PG1930	rpmC	50S ribosomal protein L29	E
			K02945	PG1297	rpsA	30S ribosomal protein S1	UA
			K02967	PG0377	rpsB	30S ribosomal protein S2	E
			K02982	PG1932	rpsC	30S ribosomal protein S3	E
			K02986	PG1912	rpsD	30S ribosomal protein S4	E
			K02988	PG1921	rpsE	30S ribosomal protein S5	E
			K02990	PG0595	rpsF	30S ribosomal protein S6	E
			K02992	PG1941	rpsG	30S ribosomal protein S7	E

			K02994	PG1924	rpsH	30S ribosomal protein S8	E
			K02946	PG1939	rpsJ	30S ribosomal protein S10	E
			K02948	PG1913	rpsK	30S ribosomal protein S11	E
			K02950	PG1942	rpsL	30S ribosomal protein S12	E
			K02954	PG1925	rpsN	30S ribosomal protein S14	E
			K02959	PG2117	rpsP	30S ribosomal protein S16	E
			K02961	PG1929	rpsQ	30S ribosomal protein S17	E
			K02963	PG0596	rpsR	30S ribosomal protein S18	E
			K02965	PG1934	rpsS	30S ribosomal protein S19	E
			K02952	PG1914	rpsM	30S ribosomal protein S13	E
			K02956	PG1758	rpsO	30S ribosomal protein S15	E
			K02996	PG0376	rpsI	30S ribosomal protein S9	E
		Aminoacyl-tRNA biosynthesis (tRNA synthetase)	K01875	PG0316	serS	seryl-tRNA synthetase	E
			K01874	PG0170	metG	methionyl-tRNA synthetase	UA
			K04567	PG1370	lysS	lysyl-tRNA synthetase	E
			K01885	PG1566	gltX	glutamyl-tRNA synthetase	E
			K01886	PG1951	glnS	glutaminyl-tRNA synthetase	E
			K01883	PG1878	cysS	cysteinyl-tRNA synthetase	E
			K01887	PG0267	argS	arginyl-tRNA synthetase	E
			K01867	PG2085	trpS	tryptophanyl-tRNA synthetase II	E
			K01889	PG1771	pheS	phenylalanyl-tRNA synthetase subunit alpha	E
			K01890	PG0099	pheT	phenylalanyl-tRNA synthetase subunit beta	E
			K01870	PG1596	ileS	isoleucyl-tRNA synthetase	E
			K01881	PG0962	proS	prolyl-tRNA synthetase	E
			K01880	PG2165	glyS	glycyl-tRNA synthetase	E
			K01872	PG1246	alaS	alanyl-tRNA synthetase	E
			K01892	PG2062	hisS	histidyl-tRNA synthetase	E
			K01873	PG1132	valS	valyl-tRNA synthetase	UA
			K01868	PG0992	thrS	threonyl-tRNA synthetase	E

			K01866	PG0263	tyrS	tyrosyl-tRNA synthetase	UA
			K01869	PG0796	leuS	leucyl-tRNA synthetase	E
			K01893	PG1121	asnC	asparaginyl-tRNA synthetase	UA
			K01876	PG0153	aspS	aspartyl-tRNA synthetase	E
			K00604	PG2023	fmt	methionyl-tRNA formyltransferase	E
		(tRNA processing)	K01056	PG0166	pth	peptidyl-tRNA hydrolase	E
			K04075	PG2046	-	hypothetical protein: tRNA(Ile)-lysidine synthase	UA
			K00554	PG2035	trmD	tRNA (guanine-N(1)-)-methyltransferase	E
			K00784	PG0739	elaC	ribonuclease Z	NE
			K03536	PG0201	rnpA	ribonuclease P protein component,	E
			K00566	PG0268	mnmA	tRNA-specific 2-thiouridylase MnmA	E
		Translation factors (Initiation factor)	K02518	PG1916	infA	translation initiation factor IF-1	E
			K02519	PG0255	infB	translation initiation factor IF-2	E
			K02520	PG0991	infC	translation initiation factor IF-3	E
		Translation factors (Elongation factors)	K02358	PG0387	tuf	elongation factor Tu	E
			K02357	PG0378	tsf	elongation factor Ts	E
			K02355	PG1940	fusA	elongation factor G	E
			K02355	PG0933	fusA	elongation factor G	NE
		Translation factors (Release factors)	K02835	PG0074	prfA	peptide chain release factor 1	E
		Translation factors (Recycling factors)	K02838	PG1901	frr	ribosome recycling factor	NE
		Protein export	K03070	PG0514	secA	preprotein translocase subunit SecA	E
			K12257	PG1762	secDF	bifunctional preprotein translocase subunit SecD/SecF	E
			K03106	PG1115	ffh	SRP54, signal recognition particle GTPase protein,	E
			K03076	PG1918	secY	preprotein translocase subunit SecY	E
			K03110	PG0151	ftsY	SRPR, signal recognition particle-docking protein,	E
		Chaperones and folding catalysts (Protein folding)	K04077	PG0520	groEL	chaperonin GroEL	E
			K04078	PG0521	groES	co-chaperonin GroES	E
			K03686	PG1776	dnaJ	chaperone protein dnaJ,	E

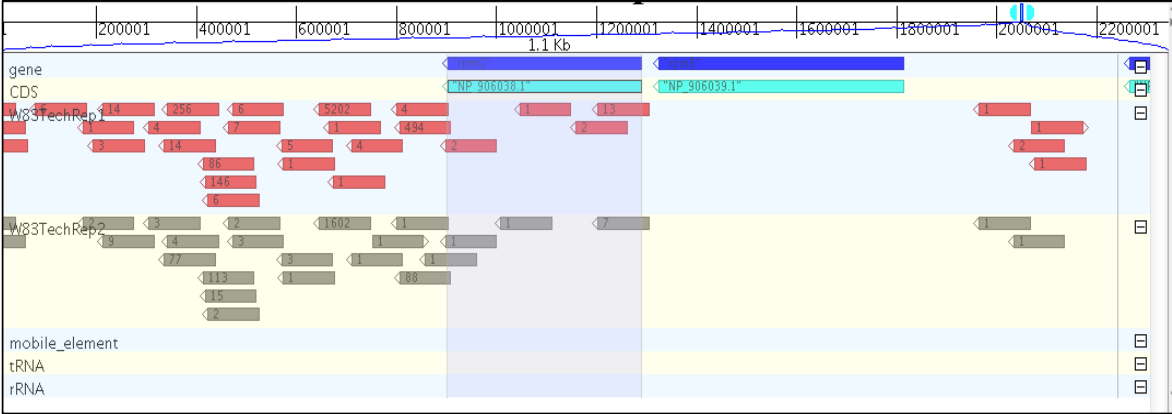
		(Protein maturation)	K04043	PG1208	dnaK	molecular chaperone DnaK	UA
			K00970	PG0801	-	poly (A) polymerase	UA
			K01265	PG1917	map	methionine aminopeptidase	E
			K01462	PG2201	def	peptide deformylase	E
	GTP-binding proteins	GTP-binding proteins	K03595	PG2142	era	GTP-binding protein Era	E
			K03979	PG0790	obgE	GTPase ObgE	E
			K03978	PG0346	engB	GTPase EngB	E
			K03977	PG2143	engA	GTP-binding protein EngA	E
	Cell division	Chromosome partitioning proteins (cell division)	K09811	PG1536	ftsX	cell division protein FtsX,	E
			K09812	PG2190	ftsE	cell division protein FtsE,	E
			K03588	PG0579	ftsW	cell division protein FtsW,	UA
			K03531	PG0584	ftsZ	cell division protein FtsZ	E
			K03590	PG0583	ftsA	cell division protein FtsA	E
			K03798	PG0047	ftsH	cell division protein FtsH	E
			K03589	PG0582	ftsQ	cell division protein FtsQ	UA
			K03569	PG1396	mreB	rod shape determining protein MreB	E
			K03570	PG1395	mreC	rod shape determining protein MreC	E
Cofactors	FMN/FAD biosynthesis	Riboflavin metabolism	K11753	PG0957	ribF	riboflavin biosynthesis protein RibF	UA
	SAM biosynthesis	Cysteine and methionine metabolism	K00789	PG1896	metK	S-adenosylmethionine synthetase	E

E = Experimentally Essential

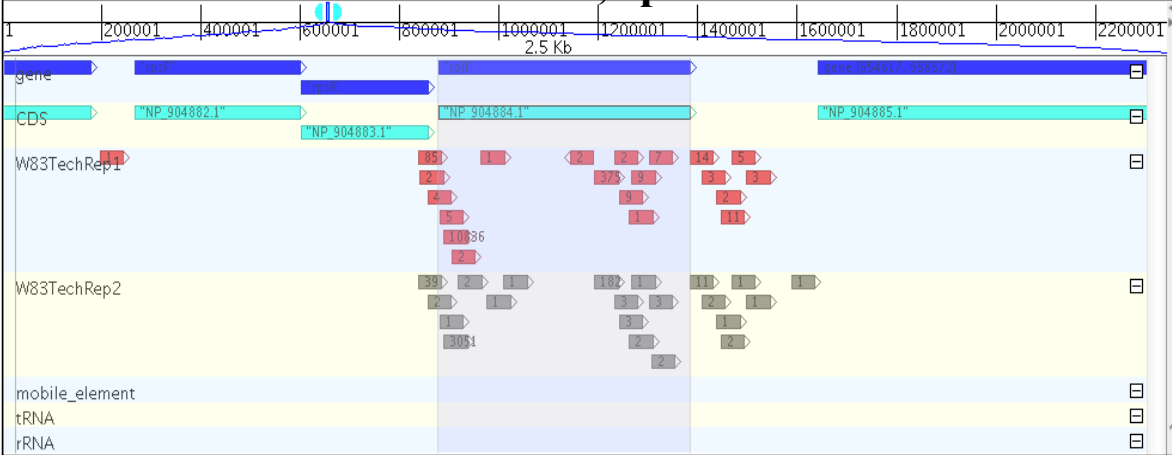
NE = Experimentally Non-essential

UA = Unaccounted in experimental data

PG1959, rpmG

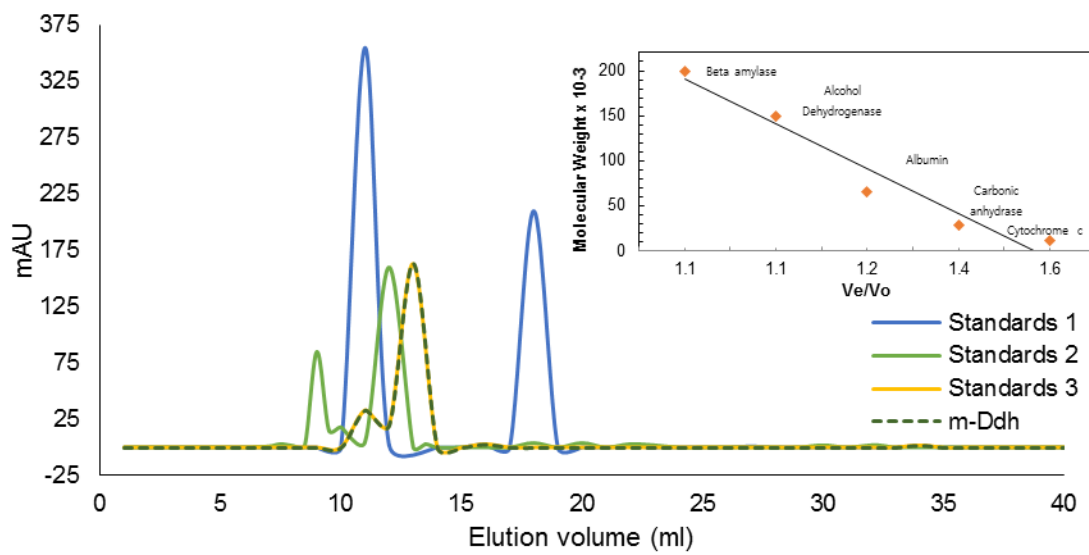
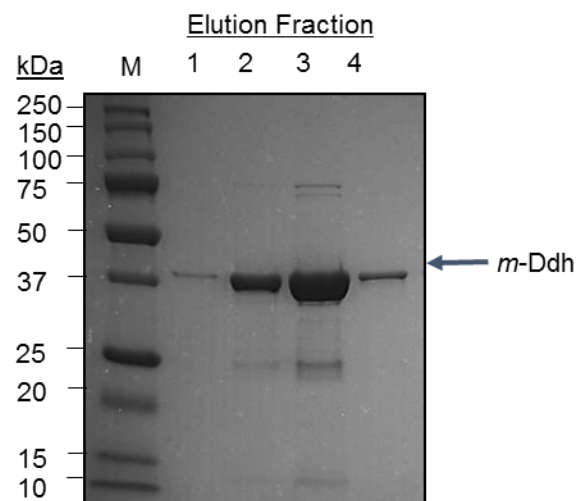


PG0392, rpIL



Supplemental Figure 1. Example of transposon insertion for *P. gingivalis* strain W83.

Top panel represents the transposon insertion pattern for an essential gene PG1959, rpmG 50S ribosomal protein L33. Bottom panel represents a transposon insertion pattern for a non-essential gene PG0392, rplL 50S ribosomal protein L10. Highlighted blue bars depict the indicated gene sequence. Red arrows represent location and orientation of a single insertion. Each gene shows insertions for two technical repeats. Courtesy of Dr. Brian A. Klein.



Supplemental Figure 2. SDS PAGE and gel filtration analysis of purified *m*-Ddh from *P. gingivalis*.

Samples were loaded on a 12.5% polyacrylamide SDS-PAGE gel and bands were detected by Coomassie G-250 stain. (M) represents the protein standards marker (Bio-Rad) and (1-4) represent collected elution fractions during purification. Arrow points to the expected size of *m*-Ddh at approximately 32 kDa. A representative gel filtration chromatograph of purified *m*-Ddh in elution buffer ran on Superdex 75 column. The native molecular mass followed with the expected peak at approximately 66 kDa similar to the protein standard albumin. Standard 1 included cytochrome c and β -amylase, Standards 2 included carbonic anhydrase and alcohol dehydrogenase and Standards 3 included albumin.

Literature Cited

1. Dewhirst FE, Chen T, Izard J, Paster BJ, Tanner AC, Yu WH, Lakshmanan A, Wade WG: **The human oral microbiome.** J Bacteriol 2010, **192**(19):5002-5017.
2. Lederberg J, McCray AT: 'Ome Sweet 'Omics - A Genealogical Treasury of Words. Scientist 2001, **15**(7):8.
3. Human Microbiome Project Consortium: **Structure, function and diversity of the healthy human microbiome.** Nature 2012, **486**(7402):207-214.
4. Wade WG: **The oral microbiome in health and disease.** Pharmacol Res 2013, **69**(1):137-143.
5. Vollaard EJ, Clasener HA: **Colonization resistance.** Antimicrob Agents Chemother 1994, **38**(3):409-414.
6. Socransky SS, Haffajee AD: **The bacterial etiology of destructive periodontal disease: current concepts.** J Periodontol 1992, **63**(4 Suppl):322-331.
7. Uehara Y, Kikuchi K, Nakamura T, Nakama H, Agematsu K, Kawakami Y, Maruchi N, Totsuka K: **H(2)O(2) produced by viridans group streptococci may contribute to inhibition of methicillin-resistant *Staphylococcus aureus* colonization of oral cavities in newborns.** Clin Infect Dis 2001, **32**(10):1408-1413.
8. Kreth J, Merritt J, Shi W, Qi F: **Competition and coexistence between *Streptococcus mutans* and *Streptococcus sanguinis* in the dental biofilm.** J Bacteriol 2005, **187**(21):7193-7203.
9. Lundberg JO, Carlstrom M, Larsen FJ, Weitzberg E: **Roles of dietary inorganic nitrate in cardiovascular health and disease.** Cardiovasc Res 2011, **89**(3):525-532.
10. Petersson J, Carlstrom M, Schreiber O, Phillipson M, Christoffersson G, Jagare A, Roos S, Jansson EA, Persson AE, Lundberg JO, Holm L: **Gastroprotective and blood pressure lowering effects of dietary nitrate are abolished by an antiseptic mouthwash.** Free Radic Biol Med 2009, **46**(8):1068-1075.
11. Jhajharia K, Parolia A, Shetty KV, Mehta LK: **Biofilm in endodontics: A review.** J Int Soc Prev Community Dent 2015, **5**(1):1-12.
12. Marsh PD: **Dental plaque as a biofilm and a microbial community - implications for health and disease.** BMC Oral Health 2006, **6** Suppl 1:S14.
13. Kolenbrander PE, Palmer RJ, Jr, Rickard AH, Jakubovics NS, Chalmers NI, Diaz PI: **Bacterial interactions and successions during plaque development.** Periodontol 2000 2006, **42**:47-79.

14. Sbordone L, Bortolaia C: **Oral microbial biofilms and plaque-related diseases: microbial communities and their role in the shift from oral health to disease.** Clin Oral Investig 2003, **7**(4):181-188.
15. Marsh PD: **Dental plaque as a microbial biofilm.** Caries Res 2004, **38**(3):204-211.
16. Loesche WJ, Gusberti F, Mettraux G, Higgins T, Syed S: **Relationship between oxygen tension and subgingival bacterial flora in untreated human periodontal pockets.** Infect Immun 1983, **42**(2):659-667.
17. Lamont RJ, Jenkinson HF: **Life below the gum line: pathogenic mechanisms of *Porphyromonas gingivalis*.** Microbiol Mol Biol Rev 1998, **62**(4):1244-1263.
18. Lamont RJ, Jenkinson HF: **Subgingival colonization by *Porphyromonas gingivalis*.** Oral Microbiol Immunol 2000, **15**(6):341-349.
19. Cekici A, Kantarci A, Hasturk H, Van Dyke TE: **Inflammatory and immune pathways in the pathogenesis of periodontal disease.** Periodontol 2000 2014, **64**(1):57-80.
20. Ebersole JL, Dawson DR, 3rd, Morford LA, Peyyala R, Miller CS, Gonzalez OA: **Periodontal disease immunology: 'double indemnity' in protecting the host.** Periodontol 2000 2013, **62**(1):163-202.
21. Ohlrich EJ, Cullinan MP, Seymour GJ: **The immunopathogenesis of periodontal disease.** Aust Dent J 2009, **54** Suppl 1:S2-10.
22. Baker PJ, Dixon M, Evans RT, Dufour L, Johnson E, Roopenian DC: **CD4(+) T cells and the proinflammatory cytokines gamma interferon and interleukin-6 contribute to alveolar bone loss in mice.** Infect Immun 1999, **67**(6):2804-2809.
23. Petersen PE: **The World Oral Health Report 2003: continuous improvement of oral health in the 21st century--the approach of the WHO Global Oral Health Programme.** Community Dent Oral Epidemiol 2003, **31** Suppl 1:3-23.
24. Batchelor P: **Is periodontal disease a public health problem?** Br Dent J 2014, **217**(8):405-409.
25. Abrahamsson KH, Wennstrom JL, Hallberg U: **Patients' views on periodontal disease; attitudes to oral health and expectancy of periodontal treatment: a qualitative interview study.** Oral Health Prev Dent 2008, **6**(3):209-216.
26. Locker D: **Oral health and quality of life.** Oral Health Prev Dent 2004, **2** Suppl 1:247-253.
27. Avila M, Ojcius DM, Yilmaz O: **The oral microbiota: living with a permanent guest.** DNA Cell Biol 2009, **28**(8):405-411.
28. Babu NC, Gomes AJ: **Systemic manifestations of oral diseases.** J Oral Maxillofac Pathol 2011, **15**(2):144-147.

29. Albandar JM: **Epidemiology and risk factors of periodontal diseases.** Dent Clin North Am 2005, **49**(3):517-32, v-vi.
30. Eke PI, Dye BA, Wei L, Slade GD, Thornton-Evans GO, Borgnakke WS, Taylor GW, Page RC, Beck JD, Genco RJ: **Update on Prevalence of Periodontitis in Adults in the United States: NHANES 2009 - 2012.** J Periodontol 2015, :1-18.
31. Eke PI, Dye BA, Wei L, Thornton-Evans GO, Genco RJ, CDC Periodontal Disease Surveillance workgroup: James Beck (University of North Carolina, Chapel Hill, USA), Gordon Douglass (Past President, American Academy of Periodontology), Roy Page (University of Washin: **Prevalence of periodontitis in adults in the United States: 2009 and 2010.** J Dent Res 2012, **91**(10):914-920.
32. Van Dyke TE, Sheilesh D: **Risk factors for periodontitis.** J Int Acad Periodontol 2005, **7**(1):3-7.
33. AlJehani YA: **Risk factors of periodontal disease: review of the literature.** Int J Dent 2014, **2014**:182513.
34. Genco RJ, Borgnakke WS: **Risk factors for periodontal disease.** Periodontol 2000 2013, **62**(1):59-94.
35. Loesche WJ: **Clinical and microbiological aspects of chemotherapeutic agents used according to the specific plaque hypothesis.** J Dent Res 1979, **58**(12):2404-2412.
36. Loesche WJ, Grossman NS: **Periodontal disease as a specific, albeit chronic, infection: diagnosis and treatment.** Clin Microbiol Rev 2001, **14**(4):727-52, table of contents.
37. Tatakis DN, Kumar PS: **Etiology and pathogenesis of periodontal diseases.** Dent Clin North Am 2005, **49**(3):491-516, v.
38. Loesche WJ: **Chemotherapy of dental plaque infections.** Oral Sci Rev 1976, **9**:65-107.
39. Rosier BT, De Jager M, Zaura E, Krom BP: **Historical and contemporary hypotheses on the development of oral diseases: are we there yet?** Front Cell Infect Microbiol 2014, **4**:92.
40. Socransky SS, Haffajee AD, Cugini MA, Smith C, Kent RL,Jr: **Microbial complexes in subgingival plaque.** J Clin Periodontol 1998, **25**(2):134-144.
41. Holt SC, Ebersole JL: ***Porphyromonas gingivalis*, *Treponema denticola*, and *Tannerella forsythia*: the 'red complex', a prototype polybacterial pathogenic consortium in periodontitis.** Periodontol 2000 2005, **38**:72-122.
42. Haffajee AD, Socransky SS: **Microbial etiological agents of destructive periodontal diseases.** Periodontol 2000 1994, **5**:78-111.
43. Nelson KE, Fleischmann RD, DeBoy RT, Paulsen IT, Fouts DE, Eisen JA, Daugherty SC, Dodson RJ, Durkin AS, Gwinn M, Haft DH, Kolonay JF, Nelson WC, Mason T, Tallon L, Gray J, Granger D, Tettelin H, Dong H, Galvin JL, Duncan MJ, Dewhirst FE,

- Fraser CM: **Complete genome sequence of the oral pathogenic Bacterium *Porphyromonas gingivalis* strain W83.** J Bacteriol 2003, **185**(18):5591-5601.
44. Bartlett JG, Gorbach SL, Finegold SM: **The bacteriology of aspiration pneumonia.** Am J Med 1974, **56**(2):202-207.
 45. Hajishengallis G, Wang M, Bagby GJ, Nelson S: **Importance of TLR2 in early innate immune response to acute pulmonary infection with *Porphyromonas gingivalis* in mice.** J Immunol 2008, **181**(6):4141-4149.
 46. Persson R, Hitti J, Verhelst R, Vaneechoutte M, Persson R, Hirschi R, Weibel M, Rothen M, Temmerman M, Paul K, Eschenbach D: **The vaginal microflora in relation to gingivitis.** BMC Infect Dis 2009, **9**:6-2334-9-6.
 47. Lamont RJ, Gil S, Demuth DR, Malamud D, Rosan B: **Molecules of *Streptococcus gordonii* that bind to *Porphyromonas gingivalis*.** Microbiology 1994, **140** (Pt 4)(Pt 4):867-872.
 48. Yao ES, Lamont RJ, Leu SP, Weinberg A: **Interbacterial binding among strains of pathogenic and commensal oral bacterial species.** Oral Microbiol Immunol 1996, **11**(1):35-41.
 49. Yang HW, Huang YF, Chou MY: **Occurrence of *Porphyromonas gingivalis* and *Tannerella forsythensis* in periodontally diseased and healthy subjects.** J Periodontol 2004, **75**(8):1077-1083.
 50. Kawada M, Yoshida A, Suzuki N, Nakano Y, Saito T, Oho T, Koga T: **Prevalence of *Porphyromonas gingivalis* in relation to periodontal status assessed by real-time PCR.** Oral Microbiol Immunol 2004, **19**(5):289-292.
 51. Fujise O, Hamachi T, Inoue K, Miura M, Maeda K: **Microbiological markers for prediction and assessment of treatment outcome following non-surgical periodontal therapy.** J Periodontol 2002, **73**(11):1253-1259.
 52. Holt SC, Kesavalu L, Walker S, Genco CA: **Virulence factors of *Porphyromonas gingivalis*.** Periodontol 2000 1999, **20**:168-238.
 53. Sheets SM, Potempa J, Travis J, Casiano CA, Fletcher HM: **Gingipains from *Porphyromonas gingivalis* W83 induce cell adhesion molecule cleavage and apoptosis in endothelial cells.** Infect Immun 2005, **73**(3):1543-1552.
 54. Olczak T, Simpson W, Liu X, Genco CA: **Iron and heme utilization in *Porphyromonas gingivalis*.** FEMS Microbiol Rev 2005, **29**(1):119-144.
 55. Chen T, Dong H, Yong R, Duncan MJ: **Pleiotropic pigmentation mutants of *Porphyromonas gingivalis*.** Microb Pathog 2000, **28**(4):235-247.
 56. Shi Y, Ratnayake DB, Okamoto K, Abe N, Yamamoto K, Nakayama K: **Genetic analyses of proteolysis, hemoglobin binding, and hemagglutination of**

- Porphyromonas gingivalis*. Construction of mutants with a combination of rgpA, rgpB, kgp, and hagA.** J Biol Chem 1999, **274**(25):17955-17960.
57. Popadiak K, Potempa J, Riesbeck K, Blom AM: **Biphasic effect of gingipains from *Porphyromonas gingivalis* on the human complement system.** J Immunol 2007, **178**(11):7242-7250.
 58. Hajishengallis G, Liang S, Payne MA, Hashim A, Jotwani R, Eskandari MA, McIntosh ML, Alsam A, Kirkwood KL, Lambris JD, Darveau RP, Curtis MA: **Low-abundance biofilm species orchestrates inflammatory periodontal disease through the commensal microbiota and complement.** Cell Host Microbe 2011, **10**(5):497-506.
 59. Isogai H, Isogai E, Yoshimura F, Suzuki T, Kagota W, Takano K: **Specific inhibition of adherence of an oral strain of *Bacteroides gingivalis* 381 to epithelial cells by monoclonal antibodies against the bacterial fimbriae.** Arch Oral Biol 1988, **33**(7):479-485.
 60. Sojar HT, Sharma A, Genco RJ: ***Porphyromonas gingivalis* fimbriae bind to cytokeratin of epithelial cells.** Infect Immun 2002, **70**(1):96-101.
 61. Evans RT, Klausen B, Sojar HT, Bedi GS, Sfintescu C, Ramamurthy NS, Golub LM, Genco RJ: **Immunization with *Porphyromonas (Bacteroides) gingivalis* fimbriae protects against periodontal destruction.** Infect Immun 1992, **60**(7):2926-2935.
 62. Song H, Belanger M, Whitlock J, Kozarov E, Progulske-Fox A: **Hemagglutinin B is involved in the adherence of *Porphyromonas gingivalis* to human coronary artery endothelial cells.** Infect Immun 2005, **73**(11):7267-7273.
 63. Katz J, Black KP, Michalek SM: **Host responses to recombinant hemagglutinin B of *Porphyromonas gingivalis* in an experimental rat model.** Infect Immun 1999, **67**(9):4352-4359.
 64. Grenier D, Mayrand D: **Functional characterization of extracellular vesicles produced by *Bacteroides gingivalis*.** Infect Immun 1987, **55**(1):111-117.
 65. Darveau RP, Pham TT, Lemley K, Reife RA, Bainbridge BW, Coats SR, Howald WN, Way SS, Hajjar AM: ***Porphyromonas gingivalis* lipopolysaccharide contains multiple lipid A species that functionally interact with both toll-like receptors 2 and 4.** Infect Immun 2004, **72**(9):5041-5051.
 66. Schwartz J, Stinson FL, Parker RB: **The passage of tritiated bacterial endotoxin across intact gingival crevicular epithelium.** J Periodontol 1972, **43**(5):270-276.
 67. Darveau RP, Arbabi S, Garcia I, Bainbridge B, Maier RV: ***Porphyromonas gingivalis* lipopolysaccharide is both agonist and antagonist for p38 mitogen-activated protein kinase activation.** Infect Immun 2002, **70**(4):1867-1873.
 68. Kumada H, Haishima Y, Umemoto T, Tanamoto K: **Structural study on the free lipid A isolated from lipopolysaccharide of *Porphyromonas gingivalis*.** J Bacteriol 1995, **177**(8):2098-2106.

69. Darveau RP, Cunningham MD, Bailey T, Seachord C, Ratcliffe K, Bainbridge B, Dietsch M, Page RC, Aruffo A: **Ability of bacteria associated with chronic inflammatory disease to stimulate E-selectin expression and promote neutrophil adhesion.** Infect Immun 1995, **63**(4):1311-1317.
70. Hajishengallis G, Darveau RP, Curtis MA: **The keystone-pathogen hypothesis.** Nat Rev Microbiol 2012, **10**(10):717-725.
71. Bik EM, Long CD, Armitage GC, Loomer P, Emerson J, Mongodin EF, Nelson KE, Gill SR, Fraser-Liggett CM, Relman DA: **Bacterial diversity in the oral cavity of 10 healthy individuals.** ISME J 2010, **4**(8):962-974.
72. Bainbridge B, Verma RK, Eastman C, Yehia B, Rivera M, Moffatt C, Bhattacharyya I, Lamont RJ, Kesavalu L: **Role of *Porphyromonas gingivalis* phosphoserine phosphatase enzyme SerB in inflammation, immune response, and induction of alveolar bone resorption in rats.** Infect Immun 2010, **78**(11):4560-4569.
73. Darveau RP, Belton CM, Reife RA, Lamont RJ: **Local Chemokine Paralysis, a Novel Pathogenic Mechanism for *Porphyromonas gingivalis*.** Infect Immun 1998, **66**:1660-1665.
74. Jagels MA, Travis J, Potempa J, Pike R, Hugli TE: **Proteolytic inactivation of the leukocyte C5a receptor by proteinases derived from *Porphyromonas gingivalis*.** Infect Immun 1996, **64**(6):1984-1991.
75. Maekawa T, Abe T, Hajishengallis E, Hosur KB, DeAngelis RA, Ricklin D, Lambris JD, Hajishengallis G: **Genetic and intervention studies implicating complement C3 as a major target for the treatment of periodontitis.** J Immunol 2014, **192**(12):6020-6027.
76. Wang M, Krauss JL, Domon H, Hosur KB, Liang S, Magotti P, Triantafilou M, Triantafilou K, Lambris JD, Hajishengallis G: **Microbial hijacking of complement-toll-like receptor crosstalk.** Sci Signal 2010, **3**(109):ra11.
77. Liang S, Krauss JL, Domon H, McIntosh ML, Hosur KB, Qu H, Li F, Tzekou A, Lambris JD, Hajishengallis G: **The C5a receptor impairs IL-12-dependent clearance of *Porphyromonas gingivalis* and is required for induction of periodontal bone loss.** J Immunol 2011, **186**(2):869-877.
78. Hajishengallis G, Wang M, Liang S, Triantafilou M, Triantafilou K: **Pathogen induction of CXCR4/TLR2 cross-talk impairs host defense function.** Proc Natl Acad Sci U S A 2008, **105**(36):13532-13537.
79. Gonzalez D, Tzianabos AO, Genco CA, Gibson FC, 3rd: **Immunization with *Porphyromonas gingivalis* capsular polysaccharide prevents *P. gingivalis*-elicited oral bone loss in a murine model.** Infect Immun 2003, **71**(4):2283-2287.
80. Hamada N, Watanabe K, Tahara T, Nakazawa K, Ishida I, Shibata Y, Kobayashi T, Yoshie H, Abiko Y, Umemoto T: **The r40-kDa outer membrane protein human**

- monoclonal antibody protects against *Porphyromonas gingivalis*-induced bone loss in rats.** J Periodontol 2007, **78**(5):933-939.
81. Maeba S, Otake S, Namikoshi J, Shibata Y, Hayakawa M, Abiko Y, Yamamoto M: **Transcutaneous immunization with a 40-kDa outer membrane protein of *Porphyromonas gingivalis* induces specific antibodies which inhibit coaggregation by *P. gingivalis*.** Vaccine 2005, **23**(19):2513-2521.
 82. Namikoshi J, Otake S, Maeba S, Hayakawa M, Abiko Y, Yamamoto M: **Specific antibodies induced by nasally administered 40-kDa outer membrane protein of *Porphyromonas gingivalis* inhibits coaggregation activity of *P. gingivalis*.** Vaccine 2003, **22**(2):250-256.
 83. Ross BC, Czajkowski L, Hocking D, Margetts M, Webb E, Rothel L, Patterson M, Agius C, Camuglia S, Reynolds E, Littlejohn T, Gaeta B, Ng A, Kuczek ES, Mattick JS, Gearing D, Barr IG: **Identification of vaccine candidate antigens from a genomic analysis of *Porphyromonas gingivalis*.** Vaccine 2001, **19**(30):4135-4142.
 84. Darveau RP: **Periodontitis: a polymicrobial disruption of host homeostasis.** Nature Rev Microbiol 2010, **8**:481-490.
 85. Rajapakse PS, O'Brien-Simpson NM, Slakeski N, Hoffmann B, Reynolds EC: **Immunization with the RgpA-Kgp proteinase-adhesin complexes of *Porphyromonas gingivalis* protects against periodontal bone loss in the rat periodontitis model.** Infect Immun 2002, **70**(5):2480-2486.
 86. Darveau RP, Hajishengallis G, Curtis MA: ***Porphyromonas gingivalis* as a potential community activist for disease.** J Dent Res 2012, **91**(9):816-820.
 87. Page RC, Lantz MS, Darveau R, Jeffcoat M, Mancl L, Houston L, Braham P, Persson GR: **Immunization of *Macaca fascicularis* against experimental periodontitis using a vaccine containing cysteine proteases purified from *Porphyromonas gingivalis*.** Oral Microbiol Immunol 2007, **22**(3):162-168.
 88. Booth V, Ashley FP, Lehner T: **Passive immunization with monoclonal antibodies against *Porphyromonas gingivalis* in patients with periodontitis.** Infect Immun 1996, **64**(2):422-427.
 89. Col NF, O'Connor RW: **Estimating worldwide current antibiotic usage: report of Task Force 1.** Rev Infect Dis 1987, **9 Suppl 3**:S232-43.
 90. Aminov RI: **A brief history of the antibiotic era: lessons learned and challenges for the future.** Front Microbiol 2010, **1**:134.
 91. Davies J: **Origins and evolution of antibiotic resistance.** Microbiologia 1996, **12**(1):9-16.
 92. Davies J, Davies D: **Origins and evolution of antibiotic resistance.** Microbiol Mol Biol Rev 2010, **74**(3):417-433.

93. Butler MS, Blaskovich MA, Cooper MA: **Antibiotics in the clinical pipeline in 2013.** J Antibiot (Tokyo) 2013, **66**(10):571-591.
94. Powers JH: **Antimicrobial drug development--the past, the present, and the future.** Clin Microbiol Infect 2004, **10 Suppl 4**:23-31.
95. Fischbach MA, Walsh CT: **Antibiotics for emerging pathogens.** Science 2009, **325**(5944):1089-1093.
96. Projan SJ, Shlaes DM: **Antibacterial drug discovery: is it all downhill from here?** Clin Microbiol Infect 2004, **10 Suppl 4**:18-22.
97. Brotz-Oesterhelt H, Sass P: **Postgenomic strategies in antibacterial drug discovery.** Future Microbiol 2010, **5**(10):1553-1579.
98. Rossolini GM, Thaller MC: **Coping with antibiotic resistance: contributions from genomics.** Genome Med 2010, **2**(2):15.
99. Haselbeck R, Wall D, Jiang B, Ketela T, Zyskind J, Bussey H, Foulkes JG, Roemer T: **Comprehensive essential gene identification as a platform for novel anti-infective drug discovery.** Curr Pharm Des 2002, **8**(13):1155-1172.
100. Tenover FC: **Mechanisms of antimicrobial resistance in bacteria.** Am J Med 2006, **119**(6 Suppl 1):S3-10; discussion S62-70.
101. Wright GD: **Mechanisms of resistance to antibiotics.** Curr Opin Chem Biol 2003, **7**(5):563-569.
102. Centers for Disease Control and Prevention: **Antibiotic Resistance Threats in the United States, 2013.** 2013.
103. Gorwitz RJ, Kruszon-Moran D, McAllister SK, McQuillan G, McDougal LK, Fosheim GE, Jensen BJ, Killgore G, Tenover FC, Kuehnert MJ: **Changes in the prevalence of nasal colonization with *Staphylococcus aureus* in the United States, 2001-2004.** J Infect Dis 2008, **197**(9):1226-1234.
104. Chambers HF, Deleo FR: **Waves of resistance: *Staphylococcus aureus* in the antibiotic era.** Nat Rev Microbiol 2009, **7**(9):629-641.
105. BARBER M, ROZWADOWSKA-DOWZENKO M: **Infection by penicillin-resistant staphylococci.** Lancet 1948, **2**(6530):641-644.
106. BARBER M: **Methicillin-Resistant Staphylococci and Hospital Infection.** Postgrad Med J 1964, **40**:SUPPL:178-81.
107. Peacock JE,Jr, Marsik FJ, Wenzel RP: **Methicillin-resistant *Staphylococcus aureus*: introduction and spread within a hospital.** Ann Intern Med 1980, **93**(4):526-532.
108. Jacobson KR, Tierney DB, Jeon CY, Mitnick CD, Murray MB: **Treatment outcomes among patients with extensively drug-resistant tuberculosis: systematic review and meta-analysis.** Clin Infect Dis 2010, **51**(1):6-14.

109. World Health Organization: **Global tuberculosis control: surveillance, planning, financing.** *WHO report 2008* 2008, .
110. Round JL, Mazmanian SK: **The gut microbiota shapes intestinal immune responses during health and disease.** *Nat Rev Immunol* 2009, **9**(5):313-323.
111. Macpherson AJ, Harris NL: **Interactions between commensal intestinal bacteria and the immune system.** *Nat Rev Immunol* 2004, **4**(6):478-485.
112. Mazmanian SK, Kasper DL: **The love-hate relationship between bacterial polysaccharides and the host immune system.** *Nat Rev Immunol* 2006, **6**(11):849-858.
113. Mayer EA: **Gut feelings: the emerging biology of gut-brain communication.** *Nat Rev Neurosci* 2011, **12**(8):453-466.
114. Mulle JG, Sharp WG, Cubells JF: **The gut microbiome: a new frontier in autism research.** *Curr Psychiatry Rep* 2013, **15**(2):337-012-0337-0.
115. Cho I, Blaser MJ: **The human microbiome: at the interface of health and disease.** *Nat Rev Genet* 2012, **13**(4):260-270.
116. Lee YK, Mazmanian SK: **Has the microbiota played a critical role in the evolution of the adaptive immune system?** *Science* 2010, **330**(6012):1768-1773.
117. Khanna S, Pardi DS: ***Clostridium difficile* infection: new insights into management.** *Mayo Clin Proc* 2012, **87**(11):1106-1117.
118. Ley RE, Backhed F, Turnbaugh P, Lozupone CA, Knight RD, Gordon JI: **Obesity alters gut microbial ecology.** *Proc Natl Acad Sci U S A* 2005, **102**(31):11070-11075.
119. Ridaura VK, Faith JJ, Rey FE, Cheng J, Duncan AE, Kau AL, Griffin NW, Lombard V, Henrissat B, Bain JR, Muehlbauer MJ, Ilkayeva O, Semenkovich CF, Funai K, Hayashi DK, Lyle BJ, Martini MC, Ursell LK, Clemente JC, Van Treuren W, Walters WA, Knight R, Newgard CB, Heath AC, Gordon JI: **Gut microbiota from twins discordant for obesity modulate metabolism in mice.** *Science* 2013, **341**(6150):1241214.
120. Ajslev TA, Andersen CS, Gamborg M, Sorensen TI, Jess T: **Childhood overweight after establishment of the gut microbiota: the role of delivery mode, pre-pregnancy weight and early administration of antibiotics.** *Int J Obes (Lond)* 2011, **35**(4):522-529.
121. Fouhy F, Guinane CM, Hussey S, Wall R, Ryan CA, Dempsey EM, Murphy B, Ross RP, Fitzgerald GF, Stanton C, Cotter PD: **High-throughput sequencing reveals the incomplete, short-term recovery of infant gut microbiota following parenteral antibiotic treatment with ampicillin and gentamicin.** *Antimicrob Agents Chemother* 2012, **56**(11):5811-5820.
122. Hughes JP, Rees S, Kalindjian SB, Philpott KL: **Principles of early drug discovery.** *Br J Pharmacol* 2011, **162**(6):1239-1249.

123. Livermore DM, British Society for Antimicrobial Chemotherapy Working Party on The Urgent Need: Regenerating Antibacterial Drug Discovery and Development: **Discovery research: the scientific challenge of finding new antibiotics**. J Antimicrob Chemother 2011, **66**(9):1941-1944.
124. Jabes D: **The antibiotic R&D pipeline: an update**. Curr Opin Microbiol 2011, **14**(5):564-569.
125. Fleischmann RD, Adams MD, White O, Clayton RA, Kirkness EF, Kerlavage AR, Bult CJ, Tomb JF, Dougherty BA, Merrick JM: **Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd**. Science 1995, **269**(5223):496-512.
126. Payne DJ, Gwynn MN, Holmes DJ, Pompliano DL: **Drugs for bad bugs: confronting the challenges of antibacterial discovery**. Nat Rev Drug Discov 2007, **6**(1):29-40.
127. Dougherty TJ, Barrett JF, Pucci MJ: **Microbial genomics and novel antibiotic discovery: new technology to search for new drugs**. Curr Pharm Des 2002, **8**(13):1119-1135.
128. Corey R, Naderer OJ, O'Riordan WD, Dumont E, Jones LS, Kurtinecz M, Zhu JZ: **Safety, tolerability, and efficacy of GSK1322322 in the treatment of acute bacterial skin and skin structure infections**. Antimicrob Agents Chemother 2014, **58**(11):6518-6527.
129. Schmid MB: **Seeing is believing: the impact of structural genomics on antimicrobial drug discovery**. Nat Rev Microbiol 2004, **2**(9):739-746.
130. Werner E: **Systems biology: the new darling of drug discovery?** Drug Discov Today 2002, **7**(18):947-949.
131. Anderson AC: **The process of structure-based drug design**. Chem Biol 2003, **10**(9):787-797.
132. Oefner C, Parisi S, Schulz H, Lociuoro S, Dale GE: **Inhibitory properties and X-ray crystallographic study of the binding of AR-101, AR-102 and iclaprim in ternary complexes with NADPH and dihydrofolate reductase from *Staphylococcus aureus***. Acta Crystallogr D Biol Crystallogr 2009, **65**(Pt 8):751-757.
133. Prakasam A, Elavarasu SS, Natarajan RK: **Antibiotics in the management of aggressive periodontitis**. J Pharm Bioallied Sci 2012, **4**(Suppl 2):S252-5.
134. Slots J: **Selection of antimicrobial agents in periodontal therapy**. J Periodontal Res 2002, **37**(5):389-398.
135. Dar-Odeh NS, Abu-Hammad OA, Al-Omiri MK, Khraisat AS, Shehabi AA: **Antibiotic prescribing practices by dentists: a review**. Ther Clin Risk Manag 2010, **6**:301-306.
136. Cheng J, Wu W, Zhang Y, Li X, Jiang X, Wei G, Tao S: **A new computational strategy for predicting essential genes**. BMC Genomics 2013, **14**(1):910.

137. Koonin EV: **Comparative genomics, minimal gene-sets and the last universal common ancestor.** Nat Rev Microbiol 2003, **1**(2):127-136.
138. Glass JI, Assad-Garcia N, Alperovich N, Yooseph S, Lewis MR, Maruf M, Hutchison CA, 3rd, Smith HO, Venter JC: **Essential genes of a minimal bacterium.** Proc Natl Acad Sci U S A 2006, **103**(2):425-430.
139. Xu P, Ge X, Chen L, Wang X, Dou Y, Xu JZ, Patel JR, Stone V, Trinh M, Evans K, Kitten T, Bonchev D, Buck GA: **Genome-wide essential gene identification in *Streptococcus sanguinis*.** Sci Rep 2011, **1**:125.
140. Acevedo-Rocha CG, Fang G, Schmidt M, Ussery DW, Danchin A: **From essential to persistent genes: a functional approach to constructing synthetic life.** Trends Genet 2013, **29**(5):273-279.
141. Hu W, Sillaots S, Lemieux S, Davison J, Kauffman S, Breton A, Linteau A, Xin C, Bowman J, Becker J, Jiang B, Roemer T: **Essential gene identification and drug target prioritization in *Aspergillus fumigatus*.** PLoS Pathog 2007, **3**(3):e24.
142. Juhas M, Eberl L, Glass JI: **Essence of life: essential genes of minimal genomes.** Trends Cell Biol 2011, **21**(10):562-568.
143. Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H: **Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection.** Mol Syst Biol 2006, **2**:2006.0008.
144. de Berardinis V, Vallenet D, Castelli V, Besnard M, Pinet A, Cruaud C, Samair S, Lechaplais C, Gyapay G, Richez C, Durot M, Kreimeyer A, Le Fevre F, Schachter V, Pezo V, Doring V, Scarpelli C, Medigue C, Cohen GN, Marliere P, Salanoubat M, Weissenbach J: **A complete collection of single-gene deletion mutants of *Acinetobacter baylyi* ADP1.** Mol Syst Biol 2008, **4**:174.
145. Kobayashi K, Ehrlich SD, Albertini A, Amati G, Andersen KK, Arnaud M, Asai K, Ashikaga S, Aymerich S, Bessieres P, Boland F, Brignell SC, Bron S, Bunai K, Chapuis J, Christiansen LC, Danchin A, Debarbouille M, Dervyn E, Deuerling E, Devine K, Devine SK, Dreesen O, Errington J, Fillinger S, Foster SJ, Fujita Y, Galizzi A, Gardan R, Eschevins C, Fukushima T, Haga K, Harwood CR, Hecker M, Hosoya D, Hullo MF, Kakeshita H, Karamata D, Kasahara Y, Kawamura F, Koga K, Koski P, Kuwana R, Imamura D, Ishimaru M, Ishikawa S, Ishio I, Le Coq D, Masson A, Mauel C, Meima R, Mellado RP, Moir A, Moriya S, Nagakawa E, Nanamiya H, Nakai S, Nygaard P, Ogura M, Ohanan T, O'Reilly M, O'Rourke M, Pragai Z, Pooley HM, Rapoport G, Rawlins JP, Rivas LA, Rivolta C, Sadaie A, Sadaie Y, Sarvas M, Sato T, Saxild HH, Scanlan E, Schumann W, Seegers JF, Sekiguchi J, Sekowska A, Seror SJ, Simon M, Stragier P, Studer R, Takamatsu H, Tanaka T, Takeuchi M, Thomaidis HB, Vagner V, van Dijl JM, Watabe K, Wipat A, Yamamoto H, Yamamoto M, Yamamoto Y, Yamane K, Yata K, Yoshida K, Yoshikawa H, Zuber U, Ogasawara N: **Essential *Bacillus subtilis* genes.** Proc Natl Acad Sci U S A 2003, **100**(8):4678-4683.

146. Song JH, Ko KS, Lee JY, Baek JY, Oh WS, Yoon HS, Jeong JY, Chun J: **Identification of essential genes in *Streptococcus pneumoniae* by allelic replacement mutagenesis.** Mol Cells 2005, **19**(3):365-374.
147. Thanassi JA, Hartman-Neumann SL, Dougherty TJ, Dougherty BA, Pucci MJ: **Identification of 113 conserved essential genes using a high-throughput gene disruption system in *Streptococcus pneumoniae*.** Nucleic Acids Res 2002, **30**(14):3152-3162.
148. Cameron DE, Urbach JM, Mekalanos JJ: **A defined transposon mutant library and its use in identifying motility genes in *Vibrio cholerae*.** Proc Natl Acad Sci U S A 2008, **105**(25):8736-8741.
149. Klein BA, Tenorio EL, Lazinski DW, Camilli A, Duncan MJ, Hu LT: **Identification of essential genes of the periodontal pathogen *Porphyromonas gingivalis*.** BMC Genomics 2012, **13**:578-2164-13-578.
150. Moule MG, Hemsley CM, Seet Q, Guerra-Assuncao JA, Lim J, Sarkar-Tyson M, Clark TG, Tan PB, Titball RW, Cuccui J, Wren BW: **Genome-wide saturation mutagenesis of *Burkholderia pseudomallei* K96243 predicts essential genes and novel targets for antimicrobial development.** MBio 2014, **5**(1):e00926-13.
151. Veeranagouda Y, Husain F, Tenorio EL, Wexler HM: **Identification of genes required for the survival of *B. fragilis* using massive parallel sequencing of a saturated transposon mutant library.** BMC Genomics 2014, **15**:429-2164-15-429.
152. Forsyth RA, Haselbeck RJ, Ohlsen KL, Yamamoto RT, Xu H, Trawick JD, Wall D, Wang L, Brown-Driver V, Froelich JM, C KG, King P, McCarthy M, Malone C, Misiner B, Robbins D, Tan Z, Zhu Zy ZY, Carr G, Mosca DA, Zamudio C, Foulkes JG, Zyskind JW: **A genome-wide strategy for the identification of essential genes in *Staphylococcus aureus*.** Mol Microbiol 2002, **43**(6):1387-1400.
153. Ji Y, Zhang B, Van SF, Horn, Warren P, Woodnutt G, Burnham MK, Rosenberg M: **Identification of critical staphylococcal genes using conditional phenotypes generated by antisense RNA.** Science 2001, **293**(5538):2266-2269.
154. Gustafson AM, Snitkin ES, Parker SC, DeLisi C, Kasif S: **Towards the identification of essential genes using targeted genome sequencing and comparative analysis.** BMC Genomics 2006, **7**:265.
155. Luo H, Lin Y, Gao F, Zhang CT, Zhang R: **DEG 10, an update of the database of essential genes that includes both protein-coding genes and noncoding genomic elements.** Nucleic Acids Res 2014, **42**(Database issue):D574-80.
156. Roemer T, Jiang B, Davison J, Ketela T, Veillette K, Breton A, Tandia F, Linteau A, Sillaots S, Marta C, Martel N, Veronneau S, Lemieux S, Kauffman S, Becker J, Storms R, Boone C, Bussey H: **Large-scale essential gene identification in *Candida albicans* and applications to antifungal drug discovery.** Mol Microbiol 2003, **50**(1):167-181.

157. Gil R, Silva FJ, Pereto J, Moya A: **Determination of the core of a minimal bacterial gene set.** Microbiol Mol Biol Rev 2004, **68**(3):518-37, table of contents.
158. Feist AM, Palsson BO: **The growing scope of applications of genome-scale metabolic reconstructions using *Escherichia coli*.** Nat Biotechnol 2008, **26**(6):659-667.
159. Mazumdar V, Snitkin ES, Amar S, Segre D: **Metabolic network model of a human oral pathogen.** J Bacteriol 2009, **191**(1):74-90.
160. Fletcher HM, Schenkein HA, Morgan RM, Bailey KA, Berry CR, Macrina FL: **Virulence of a *Porphyromonas gingivalis* W83 mutant defective in the prrH gene.** Infect Immun 1995, **63**:1521-1528.
161. Kanehisa M, Goto S: **KEGG: Kyoto Encyclopedia of Genes and Genomes.** Nucleic Acids Res 2000, **28**:27-30.
162. Tatusov RL, Galperin MY, Natale DA, Koonin EV: **The COG database: a tool for genome-scale analysis of protein functions and evolution.** Nucleic Acids Res 2000, **28**(1):33-36.
163. Zhang R, Ou HY, Zhang CT: **DEG: a database of essential genes.** Nucleic Acids Res 2004, **32**(Database issue):D271-2.
164. UniProt Consortium: **Update on activities at the Universal Protein Resource (UniProt) in 2013.** Nucleic Acids Res 2013, **41**(Database issue):D43-7.
165. Li KB: **ClustalW-MPI: ClustalW analysis using distributed and parallel computing.** Bioinformatics 2003, **19**(12):1585-1586.
166. Robert X, Gouet P: **Deciphering key features in protein structures with the new ENDscript server.** Nucleic Acids Res 2014, **42**(Web Server issue):W320-4.
167. Vanterpool E, Roy F, Fletcher HM: **The vimE gene downstream of vimA is independently expressed and is involved in modulating proteolytic activity in *Porphyromonas gingivalis* W83.** Infect Immun 2004, **72**:5555-5564.
168. Fang G, Rocha E, Danchin A: **How essential are nonessential genes?** Mol Biol Evol 2005, **22**(11):2147-2156.
169. Gerdes SY, Scholle MD, Campbell JW, Balazsi G, Ravasz E, Daugherty MD, Somera AL, Kyrpides NC, Anderson I, Gelfand MS, Bhattacharya A, Kapatral V, D'Souza M, Baev MV, Grechkin Y, Mseeh F, Fonstein MY, Overbeek R, Barabasi AL, Oltvai ZN, Osterman AL: **Experimental determination and system level analysis of essential genes in *Escherichia coli* MG1655.** J Bacteriol 2003, **185**(19):5673-5684.
170. Joyce AR, Reed JL, White A, Edwards R, Osterman A, Baba T, Mori H, Lesely SA, Palsson BO, Agarwalla S: **Experimental and computational assessment of conditionally essential genes in *Escherichia coli*.** J Bacteriol 2006, **188**(23):8259-8271.

171. Maniloff J: **The minimal cell genome: "on being the right size"**. Proc Natl Acad Sci U S A 1996, **93**(19):10004-10006.
172. Chen WH, Trachana K, Lercher MJ, Bork P: **Younger genes are less likely to be essential than older genes, and duplicates are less likely to be essential than singletons of the same age**. Mol Biol Evol 2012, **29**(7):1703-1706.
173. Jordan IK, Rogozin IB, Wolf YI, Koonin EV: **Essential genes are more evolutionarily conserved than are nonessential genes in bacteria**. Genome Res 2002, **12**(6):962-968.
174. Wilson AC, Carlson SS, White TJ: **Biochemical evolution**. Annu Rev Biochem 1977, **46**:573-639.
175. Gu X: **Maximum-likelihood approach for gene family evolution under functional divergence**. Mol Biol Evol 2001, **18**(4):453-464.
176. Mendum TA, Newcombe J, Mannan AA, Kierzek AM, McFadden J: **Interrogation of global mutagenesis data with a genome scale model of *Neisseria meningitidis* to assess gene fitness in vitro and in sera**. Genome Biol 2011, **12**(12):R127-2011-12-12-r127.
177. Jacobs MA, Alwood A, Thaipisuttikul I, Spencer D, Haugen E, Ernst S, Will O, Kaul R, Raymond C, Levy R, Chun-Rong L, Guenther D, Bovee D, Olson MV, Manoil C: **Comprehensive transposon mutant library of *Pseudomonas aeruginosa***. Proc Natl Acad Sci U S A 2003, **100**(24):14339-14344.
178. Goodman AL, McNulty NP, Zhao Y, Leip D, Mitra RD, Lozupone CA, Knight R, Gordon JI: **Identifying genetic determinants needed to establish a human gut symbiont in its habitat**. Cell Host Microbe 2009, **6**(3):279-289.
179. Rosan B, Lamont RJ: **Dental plaque formation**. Microbes Infect 2000, **2**(13):1599-1607.
180. Pavelka MS,Jr, Jacobs WR,Jr: **Biosynthesis of diaminopimelate, the precursor of lysine and a component of peptidoglycan, is an essential function of *Mycobacterium smegmatis***. J Bacteriol 1996, **178**(22):6496-6507.
181. Cremer J, Treptow C, Eggeling L, Sahm H: **Regulation of Enzymes of Lysine Biosynthesis in *Corynebacterium glutamicum***. J Gen Microbiol 1988, **134**:3221-3229.
182. Gao X, Chen X, Liu W, Feng J, Wu Q, Hua L, Zhu D: **A Novel meso-Diaminopimelate Dehydrogenase from *Symbiobacterium thermophilum*. Overexpression, Characterization and Potential for D-amino Acid Synthesis**. Appl Environ Microbiol 2012, **78**:8595-8600.
183. Hudson AO, Klartag A, Gilvarg C, Dobson RC, Marques FG, Leustek T: **Dual diaminopimelate biosynthesis pathways in *Bacteroides fragilis* and *Clostridium thermocellum***. Biochim Biophys Acta 2011, **1814**(9):1162-1168.

184. Sakamoto S, Seki M, Nagata S, Misono H: **Cloning, sequencing, and expression of the *meso*-diaminopimelate dehydrogenase gene from *Bacillus sphaericus*.** J Mol Catal B: Enzym 2001, **12**:85-92.
185. White PJ: **The essential role of diaminopimelate dehydrogenase in the biosynthesis of lysine by *Bacillus sphaericus*.** J Gen Microbiol 1983, **129**:739-749.
186. Amini S, Tavazoie S: **Antibiotics and the post-genome revolution.** Curr Opin Microbiol 2011, **14**:513-518.
187. Keller TH, Pichota A, Yin Z: **A practical view of 'druggability'.** Curr Opin Chem Biol 2006, **10**(4):357-361.
188. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ: **Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings.** Adv Drug Deliv Rev 2001, **46**(1-3):3-26.
189. Bakheet TM, Doig AJ: **Properties and identification of human protein drug targets.** Bioinformatics 2009, **25**(4):451-457.
190. Scapin G, Reddy SG, Blanchard JS: **Three-dimensional structure of *meso*-diaminopimelic acid dehydrogenase from *Corynebacterium glutamicum*.** Biochem 1996, **35**:13540-13551.
191. Hovik H, Yu WH, Olsen I, Chen T: **Comprehensive transcriptome analysis of the periodontopathogenic bacterium *Porphyromonas gingivalis* W83.** J Bacteriol 2012, **194**(1):100-114.
192. Seringhaus M, Paccanaro A, Borneman A, Snyder M, Gerstein M: **Predicting essential genes in fungal genomes.** Genome Res 2006, **16**(9):1126-1135.
193. Chen J, Brevet A, Fromant M, Leveque F, Schmitter JM, Blanquet S, Plateau P: **Pyrophosphatase is essential for growth of *Escherichia coli*.** J Bacteriol 1990, **172**(10):5686-5689.
194. Gwynn MN, Portnoy A, Rittenhouse SF, Payne DJ: **Challenges of antibacterial discovery revisited.** Ann N Y Acad Sci 2010, **1213**:5-19.
195. DiMasi JA, Hansen RW, Grabowski HG: **The price of innovation: new estimates of drug development costs.** J Health Econ 2003, **22**(2):151-185.
196. Cain R, Narramore S, McPhillie M, Simmons K, Fishwick CW: **Applications of structure-based design to antibacterial drug discovery.** Bioorg Chem 2014, **55**:69-76.
197. Drwal MN, Griffith R: **Combination of ligand- and structure-based methods in virtual screening.** Drug Discov Today Technol 2013, **10**(3):e395-401.
198. Wilson GL, Lill MA: **Integrating structure-based and ligand-based approaches for computational drug design.** Future Med Chem 2011, **3**(6):735-750.

199. Zhang Y, Yang S, Jiao Y, Liu H, Yuan H, Lu S, Ran T, Yao S, Ke Z, Xu J, Xiong X, Chen Y, Lu T: **An integrated virtual screening approach for VEGFR-2 inhibitors.** J Chem Inf Model 2013, **53**(12):3163-3177.
200. Berman HM, Battistuz T, Bhat TN, Bluhm WF, Bourne PE, Burkhardt K, Feng Z, Gilliland GL, Iype L, Jain S, Fagan P, Marvin J, Padilla D, Ravichandran V, Schneider B, Thanki N, Weissig H, Westbrook JD, Zardecki C: **The Protein Data Bank.** Acta Crystallogr D Biol Crystallogr 2002, **58**(Pt 6 No 1):899-907.
201. Agrawal H, Kumar A, Bal NC, Siddiqi MI, Arora A: **Ligand based virtual screening and biological evaluation of inhibitors of chorismate mutase (Rv1885c) from *Mycobacterium tuberculosis* H37Rv.** Bioorg Med Chem Lett 2007, **17**(11):3053-3058.
202. Kapetanovic IM: **Computer-aided drug discovery and development (CADD): in silico-chemico-biological approach.** Chem Biol Interact 2008, **171**(2):165-176.
203. Khedkar SA, Malde AK, Coutinho EC, Srivastava S: **Pharmacophore modeling in drug discovery and development: an overview.** Med Chem 2007, **3**(2):187-197.
204. Jones G, Willett P, Glen RC, Leach AR, Taylor R: **Development and validation of a genetic algorithm for flexible docking.** J Mol Biol 1997, **267**(3):727-748.
205. Korb O, Stutzle T, Exner TE: **Empirical scoring functions for advanced protein-ligand docking with PLANTS.** J Chem Inf Model 2009, **49**(1):84-96.
206. Farzan SF, Palermo LM, Yokoyama CC, Orefice G, Fornabaio M, Sarkar A, Kellogg GE, Greengard O, Porotto M, Moscona A: **Premature activation of the paramyxovirus fusion protein before target cell attachment with corruption of the viral fusion machinery.** J Biol Chem 2011, **286**(44):37945-37954.
207. Nocek B, Mulligan R, Moy S, Joachimiak A: **Crystal Structure of oxidoreductase (Gfo/Idh/MocA family member) from *Porphyromonas gingivalis* W83.** The Midwest Center for Structural Genomics (MCSG) .
208. Berman HM, Westbrook JD, Gabanyi MJ, Tao W, Shah R, Kouranov A, Schwede T, Arnold K, Kiefer F, Bordoli L, Kopp J, Podvinec M, Adams PD, Carter LG, Minor W, Nair R, La Baer J: **The protein structure initiative structural genomics knowledgebase.** Nucleic Acids Res 2009, **37**(Database issue):D365-8.
209. Sutherland A, Caplan JF, Vederas JC: **Unsaturated alpha-aminopimelic acids as potent inhibitors of *meso*-diaminopimelic (DAP) D-dehydrogenase.** Chem Commun 1999.
210. Burley SK, Joachimiak A, Montelione GT, Wilson IA: **Contributions to the NIH-NIGMS Protein Structure Initiative from the PSI Production Centers.** Structure 2008, **16**(1):5-11.
211. Gabanyi MJ, Adams PD, Arnold K, Bordoli L, Carter LG, Flippen-Andersen J, Gifford L, Haas J, Kouranov A, McLaughlin WA, Micallef DI, Minor W, Shah R, Schwede T, Tao YP, Westbrook JD, Zimmerman M, Berman HM: **The Structural Biology**

- Knowledgebase: a portal to protein structures, sequences, functions, and methods.** J Struct Funct Genomics 2011, **12**(2):45-54.
212. Hurst T: **Flexible 3D Searching: The Directed Tweak Technique.** J Chem Inf Comput Sci 1994, **34**:190-196.
213. Irwin JJ, Sterling T, Mysinger MM, Bolstad ES, Coleman RG: **ZINC: a free tool to discover chemistry for biology.** J Chem Inf Model 2012, **52**(7):1757-1768.
214. Eugene Kellogg G, Abraham DJ: **Hydrophobicity: is LogP(o/w) more than the sum of its parts?** Eur J Med Chem 2000, **35**(7-8):651-661.
215. Misono H, Togawa H, Yamamoto T, Soda K: **Meso-alpha,epsilon-diaminopimelate D-dehydrogenase: distribution and the reaction product.** J Bacteriol 1979, **137**:22-27.
216. Misono H, Soda K: **Properties of meso-alpha,epsilon-diaminopimelate D-dehydrogenase from *Bacillus sphaericus*.** J Biol Chem 1980, **255**(22):10599-10605.
217. Akita H, Fujino Y, Doi K, Ohshima T: **Highly stable meso-diaminopimelate dehydrogenase from an *Ureibacillus thermosphaericus* strain A1 isolated from a Japanese compost: purification, characterization and sequencing.** AMB Express 2011, **1**:43.
218. Wang F, Scapin G, Blanchard JS, Angeletti RH: **Substrate binding and conformational changes of *Clostridium glutamicum* diaminopimelate dehydrogenase revealed by hydrogen/deuterium exchange and electrospray mass spectrometry.** Protein Sci 1998, **7**(2):293-299.
219. Bellamacina CR: **The nicotinamide dinucleotide binding motif: a comparison of nucleotide binding proteins.** FASEB J 1996, **10**(11):1257-1269.
220. Cox RJ, Sutherland A, Vederas JC: **Bacterial diaminopimelate metabolism as a target for antibiotic design.** Bioorg Med Chem 2000, **8**(5):843-871.
221. Cox RJ: **The DAP pathway to lysine as a target for antimicrobial agents.** Nat Prod Rep 1996, **13**(1):29-43.
222. Born TL, Blanchard JS: **Structure/function studies on enzymes in the diaminopimelate pathway of bacterial cell wall biosynthesis.** Curr Opin Chem Biol 1999, **3**:607-613.
223. Bosshard HR, Marti DN, Jelesarov I: **Protein stabilization by salt bridges: concepts, experimental approaches and clarification of some misunderstandings.** J Mol Recognit 2004, **17**(1):1-16.
224. Swinney DC, Anthony J: **How were new medicines discovered?** Nat Rev Drug Discov 2011, **10**(7):507-519.
225. Seefeld MA, Miller WH, Newlander KA, Burgess WJ, DeWolf WE, Jr, Elkins PA, Head MS, Jakas DR, Janson CA, Keller PM, Manley PJ, Moore TD, Payne DJ, Pearson S,

- Polizzi BJ, Qiu X, Rittenhouse SF, Uzinskas IN, Wallis NG, Huffman WF: **Indole naphthyridinones as inhibitors of bacterial enoyl-ACP reductases FabI and FabK.** J Med Chem 2003, **46**(9):1627-1635.
226. Payne DJ, Miller WH, Berry V, Brosky J, Burgess WJ, Chen E, DeWolf Jr WE, Jr, Fosberry AP, Greenwood R, Head MS, Heerding DA, Janson CA, Jaworski DD, Keller PM, Manley PJ, Moore TD, Newlander KA, Pearson S, Polizzi BJ, Qiu X, Rittenhouse SF, Slater-Radosti C, Salyers KL, Seefeld MA, Smyth MG, Takata DT, Uzinskas IN, Vaidya K, Wallis NG, Winram SB, Yuan CC, Huffman WF: **Discovery of a novel and potent class of FabI-directed antibacterial agents.** Antimicrob Agents Chemother 2002, **46**(10):3118-3124.
 227. Sebaugh JL: **Guidelines for accurate EC50/IC50 estimation.** Pharm Stat 2011, **10**(2):128-134.
 228. Strelow J, Dewe W, Iversen PW, Brooks HB, Radding JA, McGee J, Weidner J: **Mechanism of Action assays for Enzymes.** In *Assay Guidance Manual*. Edited by Sittampalam GS, Coussens NP, Nelson H, et al. Bethesda (MD): 2004:.
 229. Silver LL: **Challenges of antibacterial discovery.** Clin Microbiol Rev 2011, **24**(1):71-109.
 230. Meredith TC, Wang H, Beaulieu P, Grundling A, Roemer T: **Harnessing the power of transposon mutagenesis for antibacterial target identification and evaluation.** Mob Genet Elements 2012, **2**(4):171-178.
 231. Freiberg C, Fischer HP, Brunner NA: **Discovering the mechanism of action of novel antibacterial agents through transcriptional profiling of conditional mutants.** Antimicrob Agents Chemother 2005, **49**(2):749-759.
 232. Huber J, Donald RG, Lee SH, Jarantow LW, Salvatore MJ, Meng X, Painter R, Onishi RH, Occi J, Dorso K, Young K, Park YW, Skwish S, Szymonifka MJ, Waddell TS, Miesel L, Phillips JW, Roemer T: **Chemical genetic identification of peptidoglycan inhibitors potentiating carbapenem activity against methicillin-resistant *Staphylococcus aureus*.** Chem Biol 2009, **16**(8):837-848.
 233. Heath RJ, Su N, Murphy CK, Rock CO: **The enoyl-[acyl-carrier-protein] reductases FabI and FabL from *Bacillus subtilis*.** J Biol Chem 2000, **275**(51):40128-40133.
 234. Jorgensen JH, Turnidge JD: **Antibacterial susceptibility tests: dilution and disk diffusion methods.** In *Manual of Clinical Microbiology*. 9th. edition. Edited by Murray PR, Baron EJ, Jorgensen JH, Landry ML, Pfaller MA. Washington, DC: American Society for Microbiology; 2007:1152-1172.
 235. Cleland WW: **Statistical analysis of enzyme kinetic data.** Methods Enzymol 1979, **63**:103-138.

236. McGovern SL, Caselli E, Grigorieff N, Shoichet BK: **A common mechanism underlying promiscuous inhibitors from virtual and high-throughput screening.** J Med Chem 2002, **45**(8):1712-1722.
237. Seidler J, McGovern SL, Doman TN, Shoichet BK: **Identification and prediction of promiscuous aggregating inhibitors among known drugs.** J Med Chem 2003, **46**(21):4477-4486.
238. Rhuland LE: **Role of α,ϵ -diaminopimelic acid in the cellular integrity of *Escherichia coli*.** J Am Chem Soc 1956.
239. Hutton CA, Southwood TJ, Turner JJ: **Inhibitors of lysine biosynthesis as antibacterial agents.** Mini Rev Med Chem 2003, **3**:115-127.
240. Paiva AM, Vanderwall DE, Blanchard JS, Kozarich JW, Williamson JM, Kelly TM: **Inhibitors of dihydrodipicolinate reductase, a key enzyme of the diaminopimelate pathway of *Mycobacterium tuberculosis*.** Biochim Biophys Acta 2001, **1545**(1-2):67-77.
241. Abbott SD, Lane-Bell P, Sidhu KP, Vederas JC: **Synthesis and Testing of Heterocyclic Analogs of Diaminopimelic Acid (DAP) as Inhibitors of DAP Dehydrogenase and DAP Epimerase.** J Am Chem Soc 1994, **116**:6513.
242. DesJarlais RL, Seibel GL, Kuntz ID, Furth PS, Alvarez JC, Ortiz de Montellano PR, DeCamp DL, Babe LM, Craik CS: **Structure-based design of nonpeptide inhibitors specific for the human immunodeficiency virus 1 protease.** Proc Natl Acad Sci U S A 1990, **87**(17):6644-6648.
243. Gleeson MP: **Generation of a set of simple, interpretable ADMET rules of thumb.** J Med Chem 2008, **51**(4):817-834.
244. Honda K: ***Porphyromonas gingivalis* sinks teeth into the oral microbiota and periodontal disease.** Cell Host & Microbe 2011, **10**:423-425.
245. Struillou X, Boutigny H, Soueidan A, Layrolle P: **Experimental animal models in periodontology: a review.** Open Dent J 2010, **4**:37-47.
246. Foster JS, Kolenbrander PE: **Development of a multispecies oral bacterial community in a saliva-conditioned flow cell.** Appl Environ Microbiol 2004, **70**(7):4340-4348.
247. Coenye T, Nelis HJ: **In vitro and in vivo model systems to study microbial biofilm formation.** J Microbiol Methods 2010, **83**(2):89-105.
248. Clardy J, Fischbach MA, Walsh CT: **New antibiotics from bacterial natural products.** Nat Biotechnol 2006, **24**(12):1541-1550.

Vitae

Victoria Nanette Stone was born in Columbia, South Carolina. She graduated from Ridge View High School, Columbia, SC in 2002. In 2006, she received her Bachelor of Science in Biochemistry from Florida State University, Tallahassee, FL. The fall of 2010, Victoria joined the laboratory of Dr. Ping Xu through the Virginia Commonwealth University, Department of Microbiology and Immunology Ph.D. program. Her accomplishments and publications are listed below.

Fellowships:

Ruth L. Kirschstein Fellowship, 2013
National Institute of Health – National Institute of Dental and Craniofacial Research

Honors and Awards:

1st Place Clinic and Research Day Poster Award, 2014
Virginia Commonwealth University, Philips Institute for Oral Health Research

Phi Kappa Phi Honor Society, 2010
Virginia Commonwealth University Chapter

Submitted Abstracts:

Inhibitors of *meso*-diaminopimelate dehydrogenase as potential antimicrobial agents against *Porphyromonas gingivalis*. **Stone V.**, Parikh H, X. Ge and P. Xu. VCU Philips Institute for Oral Health Research Clinic and Research Day. Richmond, VA. April 2015.

Inhibitors of *meso*-diaminopimelate dehydrogenase as potential antimicrobial agents against *Porphyromonas gingivalis*. **Stone V.**, Parikh H., Ge X. and Xu P. Annual Meeting of American and International Association of Dental Research. Boston, MA. March, 2015.

Target-based Discovery of Small Compound Inhibitors against Periodontal Pathogen *Porphyromonas gingivalis*. **Stone V.**, Obaidullah A., Ge X. and Xu P. VCU Philips Institute for Oral Health Research Clinic and Research Day. Richmond, VA. April 2014.

Virtual Screening to Identify Novel Antimicrobial Agents against *Porphyromonas gingivalis*. **Stone V.**, Obaidullah A., Ge X. and Xu P. Annual Meeting of American and International Association of Dental Research. Charlotte, NC. March, 2014.

A Novel Strategy in Antimicrobial Drug Discovery against Periodontal Pathogen *Porphyromonas gingivalis*. **Stone V.**, Ge X. and Xu P. Annual Meeting of the American Society for Microbiology. Denver, CO. May, 2013.

Identification of a Novel Target in *Porphyromonas gingivalis* for Controlling Periodontal Disease. **Stone V.**, Ge X. and Xu P. Watts Day. Richmond, VA. October, 2012.

Systematic Approach to Drug Discovery in *Porphyromonas gingivalis*. **Stone V.**, Ge X. and Xu P. Systems Biology Conference. Richmond, VA. May, 2011.

Publications:

Identification of Small-Molecule Inhibitors against *meso*-2,6-Diaminopimelate Dehydrogenase from *Porphyromonas gingivalis*. **Stone V.**, Parikh H.I., El-rami F., Ge X., Chen W., Zhang Y., Kellogg G.E. and Xu P. PLoS One. 2015 (accepted Oct. 2015)

The Relationship of the Lipoprotein SsaB, Manganese, and Superoxide Dismutase in *Streptococcus sanguinis* Virulence for Endocarditis. Crump K.E., Bainbridge B., Brusko S., Turner L.S., Ge X., **Stone V.**, Xu P., Kitten T. Mol Microbiol. 2014 Jun;92(6):1243-59. doi: 10.1111/mmi.12625. PMCID:PMC 4070010

Systematic study of genes influencing cellular chain length in *Streptococcus sanguinis*. Evans, K., **Stone, V.**, Chen, L., Ge, X., Xu, P. Microbiology 2014 Feb;160(Pt 2):307-15. doi: 10.1099/mic.0.071688-0 PMCID:PMC3919539

Genome-wide essential gene identification in *Streptococcus sanguinis*. Xu, P., Ge, X., Chen, L., Wang, X., Dou, Y., Xu, Patel, J.R., **Stone, V.**, Trinh, M., Evans, K., Kitten, T., Bonchev, D., Buck, G. A. Sci Rep. 2011 1, 125. doi:10.1038/srep00125 PMCID:PMC3216606